COPENHAGEN BUSINESS SCHOOL
SOLBJERG PLADS 3
DK-2000 FREDERIKSBERG
DANMARK

WWW.CBS.DK

ISSN 0906-6934

Print ISBN: 978-87-93339-18-7
Online ISBN: 978-87-93339-19-4

KNOWLEDGE DISSEMINATION BASED ON TERMINOLOGICAL ONTOLOGIES. USING EYE TRACKING TO FURTHER USER INTERFACE DESIGN

PhD Series 18-2015

Louise Pram Nielsen

KNOWLEDGE DISSEMINATION BASED ON TERMINOLOGICAL ONTOLOGIES. USING EYE TRACKING TO FURTHER USER INTERFACE DESIGN.

The PhD School of LIMAC

PhD Series 18.2015

CBS COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

Knowledge dissemination based on terminological ontologies

Using eye tracking to further user interface design

by

Louise Pram Nielsen

Supervisor: Bodil Nistrup Madsen

Submission date: April 15, 2015.

The Doctoral School of Law, Languages, Informatics,
Operations Management, Accounting and Culture (LIMAC)

Copenhagen Business School (CBS)

Louise Pram Nielsen
Knowledge dissemination based on terminological ontologies.
Using eye tracking to further user interface design.

1st edition 2015
PhD Series 18.2015

ISSN 0906-6934

Print ISBN: 978-87-93339-18-7
Online ISBN: 978-87-93339-19-4

LIMAC PhD School is a cross disciplinary PhD School connected to research communities within the areas of Languages, Law, Informatics, Operations Management, Accounting, Communication and Cultural Studies.

To my daughters

Laura & Frida

Summary

This PhD dissertation concerns domain-specific terminology. Correct use of domain-specific terms ensures concise, consistent and valid knowledge dissemination, which may be supported by term banks. However, it is crucial that target users understand the content of term banks.

This dissertation studies knowledge dissemination, based on terminological ontologies representing domain-specific terminology, and aimed at furthering the user-interface design of a term bank. Modern technology offers unlimited opportunities to meet the needs of several target groups in one term bank by offering the possibility of choosing between different presentations, in theory, providing means for knowledge transfer across different expertise levels and entry modes.

The primary focus of this PhD dissertation is an eye-tracking experiment examining target-user performance on so-called dual-entry modes representing Danish taxation terminology, to determine whether terminological ontologies should be included in the user interface. The conventional concept-oriented articles of term banks, which describe the meaning of a term by means of text, are combined with concept-oriented diagrams, which represent the meaning of a term by means of a terminological ontology displaying the underlying concepts, relations and characteristics. The latter provides the target users with overview and allows for the inference of consistent definitions, which in principle should benefit the target user.

I conduct an eye-tracking experiment, where elements of a natural user situation are replicated for a sample of 40 professional target users of a term bank. The professionals constitute primarily legal, financial or administrative personnel with advanced working tasks. The professionals are relevant because they are expected to exhibit high expertise, but also because they are an important user group for the development and promotion of the term bank. The experimental design is guided by mixed methods combining both quantitative and qualitative auxiliary data to measure participants' expertise and to gain access to the (introspective) cognitive

processing decisive for participants' performance in the experiment. It should be noted that the motivation for term variation implies multidimensional domain cultures and complex expertise of terminology users, which I measure directly using self-rating, and indirectly using representative tasks.

The analysis of direct expertise measures shows an under-estimation bias in professional target users' self-rated expertise and only one of the proposed expertise variables, exposure to discourse, is significant. The indirect expertise measures are analyzed by means of proposed representative tasks relevant to domain-specific terminology, i.e. recalling, categorizing and reading terminology, which are evaluated by expert performance indicators comprising correctness, fastness and deepness. The proposed tasks show weak expert performance suggesting further improvement of the task designs (chapter 6). In a dual-entry mode experiment, where participants are asked multiple-choice questions and provided with complementary dual entries of text and graphics, the performance indicators are used as dependent variables in a regression approach (chapter 7):

Overall, the performance models showed that domain-specific knowledge can be transferred by means of the dual-entry modes to all professional target users. Moreover, performance is improved as the experiment proceeds, suggesting learning effects and that target users adapt to the novel dual-entry modes, i.e. terminological ontologies should be part of the user interface.

However, the performance models showed no expertise effects. There may be three explanations for this result: First, reduced expertise effects arise from the dual-entry-mode design, which potentially overload the limited cognitive capacity of participants with redundant information. Second, absent expertise effects arise from the insufficient expertise measures, which do not fully capture the dimensions of expertise, or the expertise characteristics. Third, reversed expertise effects arise from the inflexibility of experts, who may be unable to adapt to the uncommon term-bank task.

This PhD dissertation contributes to terminology research, in particular to knowledge dissemination based terminological ontologies, included in the user-interface of term banks. New experiments are needed to fully capture the complex expertise of target users searching dynamic dual-entry modes. The latter should encompass search, visualization and navigation features, without disregarding the fundamental principles of terminology.

Resumé

Det overordnede emne for denne ph.d.-afhandling er terminologi, dvs. fagtermer inden for et specifikt emne. Korrekt brug af fagtermer sikrer præcis, konsistent og aktuel formidling, og for at hjælpe brugere til at finde de rette fagtermer findes en række termbanker som i tekst beskriver termers betydning (fx fagordbøger). Det er naturligvis afgørende, at brugeren rent faktisk forstår beskrivelsen og dermed er i stand til at vælge den korrekte term – på dansk såvel som på fremmedsprog.

Denne afhandling undersøger med udgangspunkt i dansk skatteterminologi, om termbankernes beskrivende tekster rent faktisk giver det optimale udbytte for brugeren, eller om de tekstbaserede beskrivelser kan kombineres med grafiske præsentationer af fagtermers betydning.

I teorien er alle ord potentielle fagtermer, men i praksis er det kun en begrænset del – formentlig kun omkring 20 pct. – af ordene i en fagtekst, som har en præcis faglig betydning. I en fagtekst vil generelle termer som fx “årrække”, “behov” og “fordeling” ikke være fagtermer, mens “energiafgift”, “forbrugsbegrænsende afgifter” og “indkomstbeskatningen” har en præcis definition på skatteområdet og dermed er fagtermer.

En traditionel termbank forklarer udelukkende betydningen af fx fagtermen “energiafgift” med ord. Dette sætter ikke fagtermens betydning ind i et større perspektiv og giver ikke brugeren overblik over fagtermernes overordnede systematik (forholdet mellem fx “energiafgift”, “punktafgift” og “kuldioxidafgift”). Dette overblik kan skabes med en såkaldt terminologisk ontologi. Ud over den tekstmæssige forklaring viser en terminologisk ontologi også fagtermers indbyrdes relationer, og gør det dermed muligt at udlede konsistente definitioner på fagtermer. Men hidtil har brugere af fagtermer kun i meget begrænset omfang haft direkte adgang til terminologiske ontologier.

For at undersøge om det kan være hensigtsmæssigt at supplere brugergrænsefladen i en termbank på skatteområdet med en terminologisk ontologi i form af en grafisk præsentation af fagtermernes indbyrdes relationer har

jeg i forbindelse med denne afhandling gennemført en række eye tracking-eksperimenter. En såkaldt remote eye tracker sporede øjenbevægelser hos forsøgsdeltagere, som blev præsenteret for konstruerede eksempler på terminologiske ontologier på skatteområdet.

Eye tracking-eksperimenterne blev gennemført med 40 forsøgsdeltagere. Alle var vant til at udføre akademiske arbejdsopgaver, og 20 af disse var medarbejdere fra SKAT. De resterende 20 var en blandet gruppe bestående af bl.a. af forskere fra CBS, jurister, økonomer, journalister og translatører. Det primære udvælgelseskriterium var, at forsøgsdeltagerne potentielt ville have en relativ høj grad af ekspertise på skatteområdet i forhold til den generelle befolkning, da det netop er personer som disse, som vil være den primære målgruppe for en termbank.

Resultaterne af eye tracking-eksperimenterne viste, at forsøgsdeltagerne havde nytte af de terminologiske ontologier, men at en høj grad af ekspertise på skatteområdet ikke var afgørende for forståelsen af de konkrete foreviste eksempler. Dette resultat kan have tre mulige forklaringer, som kan benævnes reduceret ekspertiseeffekt, fraværende ekspertiseeffekt og omvendt ekspertiseeffekt.

Reduceret ekspertiseeffekt er, at det kun giver en lille fordel at have en høj grad af ekspertise. I forsøget kan denne effekt have spillet ind, fordi eksemplernes kombination af tekst og grafik belastede forsøgsdeltagerne med for mange informationer, og forsøgsdeltagerens ekspertise dermed ikke har givet den forventede gevinst. Fraværende ekspertiseeffekt er, at ekspertise ingen rolle spiller. I forsøget kan denne effekt have spillet ind, fordi forsøgsdeltagernes ekspertise ikke er målt tilstrækkeligt præcist under forberedelsen til eksperimentet. Omvendt ekspertiseeffekt er, når det er en ulempe med høj ekspertise. I forsøget kan denne effekt have spillet ind, fordi, forsøgspersonerne kan have været belastede af deres viden og vaner. Hvis fx en forsøgsdeltager med høj ekspertise på skatteområdet præsenteres for eksperimentets definition af “energiafgift”, kan definitionens ordlyd eller relation til andre begreber adskille sig marginalt fra personens faglige forståelse. Den grafiske fremstilling i den terminologiske ontologi kan desuden virke fremmedartet for forsøgsdeltageren, og den manglende genkendelse kan føre til forvirring. Dermed slår den omvendte ekspertiseeffekt igennem.

Denne ph.d.-afhandling bidrager til forskningen i terminologi og særligt terminologiske ontologier – ikke alene i termbanker men også generelt. Forskningen rejser samtidig en række nye spørgsmål, som kun kan besvares ved nye tilpassede eksperimenter. Det er fortsat en udfordring at gøre termbanker mere brugervenlige end den traditionelle tekstbaserede opbygning tillader, og det bør yderligere undersøges om dette kan opnås ved at benytte nytænkende metoder til at søge, visualisere og navigere i termbanker – uden at gå på kompromis med de terminologiske principper.

Acknowledgments

Thank you to the VELUX FOUNDATION for funding the DanTermBank project, to the Department of International Business Communication (IBC) at CBS for hosting my doctoral studies, and to the project team members Anna Elisabeth Odgaard Ingram, Pia Lyngby Hoffmann, Radu Dudici, Tine Lassen, Hanne Erdman Thomsen and Bodil Nistrup Madsen for support. Indeed, I would like to thank my supervisors and colleagues at IBC: Bodil Nistrup Madsen, for always keeping a nice and caring spirit, even during difficult times, and especially during the final sprint. Henrik Selsøe Sørensen, for readily sharing constructive ideas and perspectives. Daniel Barrett, for getting soft- and hardware to work together during the eye-tracking experiments and for turning his office into an eye-tracking lab. Laura Winther Balling, for avoiding fishing trips in the large data pond, thereby sharpening my arguments in my statistical analyses. Rita Lenstrup, for encouragement, knowledge sharing, and for offering world-class proof reading caring for the many wonderful colours of the English language.

At SKAT, the management has been indispensable for the pursuit of my research ambitions. Firstly, I owe Tage Christensen my best thanks for supporting the application of an industrial PhD, and secondly, Hans From, for encouraging my participation in the DanTermBank project as PhD fellow. Thirdly, I thank Anne Sophie Stricker Springborg for ready assistance during maternity leave and smooth return into new structures. Moreover, I am very thankful to my participants for taking the time and effort to come to Dalgas Have and try an eye-tracking experiment, which must have felt strange to many of you.

Last, but not least, I am very thankful to my parents Poula & Mogens Pram Nielsen for looking after my young children, while I finished this dissertation. Moreover, I express my sincere gratitude to my husband, Ebbe Rasch. Together, we stayed the course and passed several major milestones, not only in academia but as parents and constructors. Thank you for your encouragement, love and commitment, and for furthering my strive for knowledge.

Ejby, April 15, 2015.

Contents

List of tables 17

List of figures 18

Abbreviations and notations 19

1 Introduction 21

1.1 Motivation 21
1.2 Research strategy 24
1.3 Outline of the dissertation 27
1.4 Relevant research areas 28

2 Domain-specific terminology 31

2.1 Introduction 31
2.2 Terminology theory and practice 37
2.2.1 Terms furthering specialized discourse 38
2.2.2 Terms constituting lexical units 42
2.2.3 Terms designating concepts 44
2.3 Term banks 48
2.3.1 Terminological ontologies 51
2.3.2 Specialized dictionaries 54
2.3.3 Combining lexicography and terminology 55
2.4 Conclusion 57

3 Terminology and knowledge 59

3.1 Introduction 60
3.2 Intra-domain cultures 63
3.2.1 Communities of practice 64
3.2.2 Examples of specialization 65
3.3 Extra-domain cultures 67
3.3.1 Communities of knowledge 68
3.3.2 Examples of fundamental concepts 70
3.4 Supra-domain cultures 71
3.4.1 Communities of language 71
3.4.2 Examples of equivalents 74
3.5 Graphic illustrations of the domain-culture framework 79

3.6 Knowledge acquisition from term banks 82
3.6.1 Declarative knowledge 83
3.6.2 Procedural knowledge 84
3.6.3 Conditional knowledge 85
3.6.4 Meta-knowledge 86
3.7 Conclusion 87

4 Experimental terminology **91**
4.1 Introduction 91
4.2 Participants 96
4.3 Self-rating tasks 99
4.4 Representative tasks 101
4.4.1 Recalling 103
4.4.2 Categorizing 103
4.4.3 Reading 105
4.4.4 Card-sorting 106
4.5 Term bank tasks 107
4.5.1 Dual-entry mode (stimulus pair) 108
4.5.2 Question types, format and feedback 110
4.5.3 Experimental procedure 115
4.6 Retrospective interview 118
4.7 Focus-group discussion 120
4.8 Conclusion 121

5 Direct expertise measures **123**
5.1 Introduction 123
5.2 Sampling across expertise variables 126
5.2.1 Participation 126
5.2.2 Motivation 126
5.2.3 Exposure 127
5.2.4 Education 127
5.3 Results: Domain-specific expertise 128
5.3.1 Raw data for self-rated expertise 128
5.3.2 Regression model for self-rated expertise 132
5.3.3 Validity of self-rated expertise 133
5.4 Results: Reference skills 134
5.4.1 Raw data for self-rated reference skills 135
5.4.2 Used and non-used reference works 136
5.4.3 Validity of self-rated reference skills 138
5.5 Conclusion 138

6 Indirect expertise measures **141**
6.1 Introduction 141
6.2 Results: Recalling terms 142
6.2.1 Correctness: Term lists 143
6.2.2 Depth: Frequency of domain 145
6.3 Results: Categorizing terms and pseudo-terms 148
6.3.1 Correctness and significant predictors 149
6.3.2 Depth: Detail of definitions 151
6.4 Results: Reading terms in an authentic, specialized text . . 151
6.4.1 Speed: Processing time of depth 152
6.4.2 Correctness: Difficulty of content 155
6.5 Results: Knowledge structure 156
6.5.1 Correctness and depth: Evaluating concept maps 157
6.5.2 Inferring expertise from mapping strategies 159
6.6 Conclusion 162

7 Dual-entry mode experiment **165**
7.1 Introduction 165
7.2 Regression approach 167
7.3 Results: Correctness 171
7.3.1 Significant predictors of correctness 172
7.3.2 Discussion of correctness 175
7.4 Results: Response time 177
7.4.1 Significant predictors response time 178
7.4.2 Discussion of response time 182
7.5 Results: Diagram fixation 182
7.5.1 Significant predictors of diagram fixation 184
7.5.2 Discussion of diagram fixation 187
7.6 Conclusion 188

8 Future research **191**
8.1 Results and implications 191
8.2 Limitations 192
8.3 Terminology and knowledge engineering issues 194
8.4 Usability engineering issues 196
8.5 Conclusion 199

Appendix **201**

Appendix I: Systematic list of concepts 202

Appendix II: Terminological ontology 213

Appendix A: Consent form and background questionnaire . . . 233

Appendix B: Introduction, recalling and categorization 240

Appendix C: Scenario and reading task 248

Appendix D: Dual-entry-mode experiment 253

Appendix E: Retrospective interview 272

Appendix F: Card-sorting 277

Appendix G: Focus-group discussions 290

Bibliography **293**

List of tables

2.1 Concept clarification 45
2.2 Terminology visualizations in selected terminological resources 50

3.1 Target-user matrix 67
3.2 Translation strategies across levels of equivalence 74
3.3 Domain-specific knowledge acquisition 84

4.1 Performance indicators and experimental tasks 102
4.2 Question types belonging to each block 111
4.3 Retrospective interview 119
4.4 Incentives to and barriers against term-bank use 121

5.1 Self-rated expertise, reference skills and English language skills 129
5.2 Summary of regression model for self-rated expertise 132
5.3 The use and knowledge of reference works 137

6.1 Recalling ten terms 143
6.2 Terms mentioned by five participants or more 144
6.3 Depth performance indicator of recalling task 146
6.4 Depth performance indicator and expertise variables 147
6.5 Terms and pseudo-terms of the categorization task 149
6.6 Summary of regression model for correctness in the categorization task 150
6.7 Processing of the reading task 153
6.8 Processing and expertise variables 154
6.9 Components of the reference map 157
6.10 Interconnectedness of the produced diagrams 158
6.11 Reference-map correspondence 159
6.12 Mapping strategies 160

7.1 Overview of regression models 170
7.2 Summary of regression model for correctness 172
7.3 Summary of regression model for response time on correct answers 178
7.4 Summary of regression model for diagram-fixation time on correct answers 183
7.5 Interpretation of diagram-fixation time 186

List of figures

2.1 The scale of specificity of linguistic units . . . 36
2.2 Domain-specific terminological ontology . . . 46

3.1 The motivation for term variation in communities of knowledge, practice and language . . . 80
3.2 Examples of domain-specific term variation . . . 81

4.1 Usability ontology . . . 97
4.2 Self-rating scale of expertise . . . 100
4.3 Example of dual-entry mode . . . 109
4.4 Visualization of eye movements . . . 116

5.1 Self-rated expertise related to motivation and participation . . . 130
5.2 Self-rated expertise related to exposure and gender . . . 131
5.3 Self-rated reference skills related to self-rated expertise and age . . . 135
5.4 Self-rated expertise and reference skills related to gender . . . 136

6.1 Depth of recalling task related to exposure and self-rated expertise . . . 148
6.2 Processing related to exposure and self-rated expertise . . . 155
6.3 Processing related to perceived text difficulty . . . 156

7.1 Predictors of correctness . . . 173
7.2 Predictors of response time on correct answers . . . 179
7.3 Random slopes of each participant . . . 181
7.4 Predictors of diagram fixation on correct answers . . . 184
7.5 Interaction between question type and trial number on diagram-fixation time . . . 187

Abbreviations and notations

AOI	Area Of Interest
A	Article condition
D	Diagram condition
DA	Diagram-Article condition
EU	The European Union
ISO	The International Organization for Standardization
LGP	Language for general purpose
LSP	Language for specialized purpose
ms	Milliseconds
n.a.	Missing or non-available information
OECD	The Organization for Economic Cooperation and Development
SKAT	the Danish Customs and Tax Administration (Denmark's tax authority)
SL	Source language or native language (L1), which in this case study is Danish.
TL	Target language is foreign language (i.e. L2, L3, etc.) but including lingua franca (regional or global English).

All translations from Danish into English are glossed, i.e. literally translated suggestions, to ease the reading. It should be noted that a thorough concept clarification may arrive at nonequivalence or near equivalence. If the point of the example is lost in translation, it is marked by ⟐*,* e.g. "act on investor funds" (*lov om investorfonds*) ⟐. Danish terms and citations are italicized and in brackets, while English translations are between double quotation marks and not italicized, e.g. "motor vehicles tax" (*afgift af motorkøretøj*) and "I've always said that taxation is beyond me!" (*Jeg har altid sagt, at skat ikke er til at forstå!*).

In the appendix, it was necessary to use a slightly different notation for the concept clarification (appendices I-II). The experimental material (appendices A-E) as well as the workshop material (appendices F-G) were presented to participants in Danish and are translated into English, where appropriate notation is used depending on the format of the content.

Chapter 1: Introduction

1.1 Motivation

My personal motivation for conducting this dissertation research is a strong terminological interest in my background field: public finance. In particular, domain-specific terminology provides means to ensure unambiguous specialized communication and knowledge dissemination. In addition, I am motivated by the challenges that modern information technology poses on the terminology and knowledge to be acquired from various forms of computerized terminology resources. Current research emphasize the development of rapid methods to provide users with big quantities of validated data as well as considering the needs and expertise of different target users (Madsen, Thomsen, Halskov & Lassen, 2010).

In addition, the technological development carries a huge potential for disseminating terminology and knowledge by combining recent lexicographic research on multimodal electronic dictionaries (Lew & de Schryver, 2014) with terminological ontology engineering (Roche, 2006). This dissertation research proposes novel approaches to the dual presentation of domain-specific terminology in text and graphics, to multiple domain-specific cultures, and to the analysis of expertise effects on target users' performance, which will further the user interface design of terminology and knowledge banks (see section 1.3 for the outline of the dissertation). The research constitutes a sub-project of the DanTermBank project, which aims at developing a national terminology and knowledge bank in Denmark (DanTermBank, 2015).

Term banks

A variety of computerized terminological resources exist including the simple term-oriented glossaries and specialized dictionaries as well as the concept-oriented and highly structured taxonomies, terminological ontologies and knowledge bases. There are various ways of referring to this type of information tool, which according to Verlinde, Leroyer & Binon (2009, p. 3) comprise “general dictionaries, specialized dictionaries, encyclopedias, terminological databases, lexical databases, glossaries, registers” in addition to universal and domain ontologies and taxonomies (Grabar, Hamon & Bodenreider, 2012). A terminology and knowledge bank can be defined as an information tool, which covers the whole spectrum (see section 2.3), as is indicated by the name containing both the “terminology"-orientation and the conceptual “knowledge”. But for the sake of my readers, I choose to use the simpler form “term bank” throughout this dissertation, despite the risk that some readers thereby mistakenly anticipate technologies merely presenting word lists instead of conceptually structured resources, where users acquire domain-specific terminology and knowledge.

User-oriented research

As the title of the dissertation suggests, “knowledge dissemination” implies transferring knowledge to potential end-users of the term bank, i.e. facilitating human knowledge acquisition, by means of a term bank to potential end-users. Consequently, the research is user-oriented. I introduce “target user” as the short term for target user groups i.e. the groups of end-users that I target with my research. It should be noted that it lies implicit that the target users are users of a term bank, but I prefer to call them target users instead of term-bank users. I abstain from the term persona (Nielsen, 2004), because that entails a different approach by the “description of a specific person who is a target user of a system being designed, providing demographic information, needs, preferences, biographical information, and a photo or illustration” (Usability First, 2015).

In research aimed at developing term banks, we are accounting for three primary types of users: The knowledge engineer (including terminologists, computational linguists and system developers) building the terminology and knowledge bank, the domain expert validating the content and representation, and the target user who acquires knowledge by using the system (Chein, Mugnier & Croitoru, 2013), though the latter two are not strictly separated. I stress that my research merely aims at furthering the interface design process, because I only include target users in my analysis, but “ideally, every stakeholder should be accounted for somewhere in the design process,” (Usability First, 2015).

I target my research to one particular user group. In my view, the most important user group to address in the development of term banks is the professionals, as they are most likely to cover the full scale of expertise from high to low, which is not necessarily the case with other user groups. Moreover, professionals possess great potential in the launch and promotion of a terminology and knowledge bank through their widespread contact with non-professional user groups in both ends of the scale of expertise.

Furthering user interface design

The user interface design process covers the “overall process of designing how a user will be able to interact with a software application,” where the user interface is defined as the “parts of a computer system that a person uses to communicate with the computer,” (Usability First, 2015).

An iterative design process constitutes “the idea that design should be done in repeated cycles where, in each cycle, the design is elaborated, refined, and tested, and the results of testing at each cycle feed into the design focus of the next cycle,” (Usability First, 2015). Throughout the dissertation, I use the term “dual-entry mode” emphasizing the “two-in-one” duality of the information formats (textual versus graphical). Strictly speaking, I only propose the design of the (static) stimuli, which are used in the eye-tracking experiments, and do not proceed into conduct any further iterations in the design process like for instance including the dual-entry modes into a realistic term-bank prototype with (dynamic) navigational functionalities. In other words, the research can be seen as a form of prototyping, from where designs should be further “created, evaluated, and refined until the desired performance or usability is achieved,” (Usability First, 2015).

Domain-specific terminology

I limit my research to the term banks containing domain-specific terminology. Compared to general-language terminology, domain-specific terminology comprises the terms and underlying concepts that are specific and belonging uniquely to a particular domain. A domain is an area of knowledge but should not be confused with a scientific discipline as a domain may comprise anything from several disciplines (e.g. public finance versus law) to sub-disciplines (e.g. revenue collection versus tax planning). If terminology can be specific to a domain, then the same terminology is necessarily non-specific to other domains. Some domains are closer than others e.g. taxation is very close to public finance (social science), while distinct from medicine (natural science). Therefore, we may evaluate terminology on a scale of specificity (see section 2.1).

The taxation domain is chosen as the domain relevant for exploring domain-specific terminology and knowledge dissemination. This choice of domain implies increased complexity due to multiple domain-specific cultures, and perhaps therefore, a rarely chosen domain in terminology research. I choose the taxation domain to explore terms empirically using semantics, pragmatics and eye-tracking experiments.

Taxation is a highly abstract topic, but we encounter taxation on a daily basis, as we carry the burden of it and pay the sales tax, excise duties, income tax, corporation tax etc. Perhaps, this is one of the reasons why the cognition and communication of tax matters are difficult and complex, because hardly any visual representations are appropriate and can be accompanied by textual documents. And yet we know (Boll, 2011) that instruction and communication with tax authorities are crucial to compliance behavior and eventually the tax collection.

1.2 Research strategy

The overall aim of this dissertation research is to further the interface design process of a term bank. The overall research question is:

How should we disseminate domain-specific terminology and knowledge?

In particular, I wish to inquire into the semantics and pragmatics of domain-specific terminology, which will provide for adaptive terminology visualization in the term-bank interface ensuring efficient knowledge dissemination to target users. I use taxation as my exploratory domain, professionals as the chosen target users, and apply eye-tracking technology in my experiments. The overall research question is elaborated into the following three underlying research questions (or themes developing initial ideas into specific hypotheses), which are motivated by the constituents of human-computer interaction (HCI):

(1) The first underlying research theme concerns **computer**ized terminology and the design of a user interface of a term bank. At first sight, it is a question of using the conventional format of term banks (see table 2.2), which constitutes concept articles. However, the concept diagram displaying the underlying terminological ontology in graph format may complement the textual articles in a dual-entry mode expanding the amount of presented knowledge and increasing the knowledge-acquisition potential.

In analyzing the first underlying research theme, I take a semantic point of view to domain-specific terms (chapter 2) to develop a term bank, which follows the formal principles of terminological ontologies (Madsen & Thomsen, 2015) as well as the user-oriented lexicographical functions (Kwary, 2012). In particular, I combine the (term-oriented) semasiologic approach with the (concept-oriented) onomasiologic approach to develop the dual-entry modes containing concept articles (text) and concept diagrams (graphics) as complementary ways to convey knowledge to target users.

(2) The second underlying research theme concerns domain-specific expertise guiding the user adaption of an interface to **human** target users. At first sight, it is a question of the conventional distinction of expertise into expert, semi-expert and layman levels. However, the direct measures do not encompass the multiple dimensions of domain-specific culture shared by communities of practice, knowledge and language.

In analyzing the second underlying research theme, I approach the users of domain-specific terminology from a pragmatic point of view (chapter 3), which allows us to capture the relevant domain cultures. I account for the three culture-specific motivations behind term variation in specialized communication and knowledge dissemination: Intra-domain cultures are shaping the terms and transferring knowledge across different levels of specialization (expert, semi-expert and layman) inside a complex domain (see e.g. Al-Sayed & Ahmad, 2006). Extra-domain cultures stem from the inter-disciplinary challenges of conveying knowledge from one discipline to another with a different set of conceptual structures (see e.g. Fernández-Silva, Freixa & Cabré, 2011). Supra-domain cultures stem from international encounters at national, regional and global levels including translation across organizational and institutional barriers crucial to the chosen domain (see e.g. Sandrini, 1999).

The framework underlines the complexity of expertise, which should be accounted for in the assessment, otherwise we will not be able to demonstrate expertise effects necessary for user adaption. I propose direct measures of expertise reflected by relevant background variables (so-called participation, motivation, education and discourse exposure) and self-rating (chapter 5), as well as indirect measures in the form of four representative tasks to domain-specific terminology (recalling, categorizing, reading and structuring) with associated performance indicators (correctness, speed and depth) supposedly demonstrating superior expert performance (chapter 6).

(3) The third underlying research theme concerns the **interaction** between target users and the dual-entry mode, which is approached by means of experimental usability testing. I test the proposed dual-entry-mode design to target users by applying eye tracking and mixed effects models providing us with so-called random effects of participants allowing us to interpret results as learning effects (implying users adapting to the design) and expertise effects (implying adaption to users).

In analyzing the third underlying research theme, I propose an experimental approach to capture target-user cognition and infer design recommendations in order to further the interface design process of a term bank: First, participants, tasks and interviews with participants are outlined (chapter 4). Then, I conduct an eye-tracking experiment using dual-entry-mode stimuli asking participants highly manipulated multiple-choice questions (chapter 7).

My research strategy combines quantitative and qualitative empirical methods (chapter 4) to gain access to informants' domain-specific expertise and (introspective) cognitive processing. The chosen regression approach to the data collected in the dual-entry experiment is necessarily quantitative, where hypotheses are formulated in terms of dependent and independent variables captured in each of the regression models. This approach is combined with semistructured qualitative data from the retrospective interview, where participants rated their view on performance and difficulty as well as preferences and professional needs pertaining to term banks. Finally, a card-sorting approach to evaluate users' understanding and structuring of the domain-specific terminology into concept maps and a focus-group discussion on term-bank use constitute qualitative data.

The combination of concept diagrams, expertise effects and eye tracking in the dissemination of Danish taxation terminology is unique: There is to my knowledge, no previous studies of cognitive processing of dual-entry modes by professionals (outlined in chapter 7) focusing on concept diagrams (discussed in chapters 2) in combination with concept articles, and considering complex domain-specific expertise as a determiner of target-user performance (discussed in chapters 3, 5 and 6).

1.3 Outline of the dissertation

This dissertation contains eight chapters followed by nine appendices and a bibliography (structured alphabetically into authors and web sites). A list of tables, figures and abbreviations are placed at pages 15-17.

After the introduction (**chapter 1**), I outline the defining criteria for domain-specific terms and discuss the terminological ontologies as potential visualization format (entry mode) of term banks (**chapter 2**). Then I turn to the target users of terminology. The motivation for term variation renders multiple domain cultures, which are crucial to the assessment of complex target-user expertise determining knowledge-acquisition potential of target users (**chapter 3**). I outline my experimental research strategy and describe the chosen participants, the tasks and interviews (**chapter 4**). The expertise of term-bank users is key to the efficiency of terminology and knowledge dissemination and often used to guide user adaption of information tools. I begin by proposing crucial expertise variables as well as direct measures asking participants to self-rate their domain-specific expertise (**chapter 5**). Moreover, I propose indirect measures by associating expert performance indicators (i.e. correctness, speed and depth) to proposed representative tasks asking participants to recall, categorize, read and structure domain-specific terminology (**chapter 6**). Then, I analyze target-user performance in the dual-entry mode experiments, where participants retrieve answers to the concept clarifying questions from the stimuli visualizing terminological data in the textual and graphical formats (**chapter 7**). The dissertation is rounded off with a discussion of the conclusions in terms of terminology and knowledge engineering and usability engineering in the light of outstanding future work (**chapter 8**).

The concept clarification of the Danish taxation terminology is presented by means of a systematic list and a terminological ontology in appendices I and II. The experimental material used in the eye-tracking work and data collection with sampled target users is translated into English and placed in appendices A-E. The workshop material used in the exercises and discussions with focus groups are presented in appendices F and G.

1.4 Relevant research areas

The aim of this dissertation research is to further the interface design of a term bank. As mentioned, the outlined research strategy (see section 1.3) combines each of the components key to the field of HCI: I develop novel dual-entry modes (i.e. terminology visualizations on the computer), which I use as stimuli in an eye-tracking experiment (i.e. to infer the cognitive processing underlying interaction), where I investigate expertise effects on performance (i.e. adapting the term bank to human users). In other words, I combine research on terminology, cognition and expertise.

Overall, the combination of numerous research fields proves a challenge to my literature review. A narrow literature search for "multimodal terminology processing" gives no results, while the broad search for literature on "terminology processing" does provide results, but not the relevant ones as the focus of this body of literature is on term extraction and machine translation not on (human) cognitive processing. In cognitive psychology, as well as the fields of HCI and usability research, eye tracking has received extensive attention. To my knowledge, however, eye-tracking technology has not been applied in terminology research. The psycholinguistic literature provides us with experimental tasks aimed at uncovering relatively low-level language processing, which only to a limited extent is applicable in our research as the pragmatic characteristics of terminology necessitates the consideration of expertise effects.

In other words, my research is not devoted to any existing paradigm in particular but rather combines elements from research on terminology, knowledge engineering, ontology engineering, specialized lexicography and computational linguistics (see chapter 2); cross-cultural communication, specialized communication and translation studies (see chapter 3); user-centered research (see chapter 4); expertise research, instructional research and educational research (see chapters 5 and 6); eye-tracking research, information retrieval, knowledge acquisition, cognitive science, experimental psychology and mixed methods (see chapter 7). The used literature is diverse and literature reviews are given chapter by chapter, focusing on those aspects particularly relevant to the underlying research themes.

Chapter 2: Domain-specific terminology

Chapter 2 is central to analyzing the first underlying research theme concerning the dual-entry mode visualizing domain-specific terminology in a term bank in concept articles and diagrams.

In this chapter, I approach terminology from a semantic point of view to explore the dual meaning of domain-specific terms comprising a lexical and a conceptual side. In particular, I combine user-oriented lexicography with the formal principles of terminological ontology to visualize terms.

I present and discuss the relevant terminology literature. The section does not constitute a general review of terminology, ontology and lexicography but rather a review of those aspects which are particularly relevant for domain-specific terminology visualized by the proposed dual-entry modes (section 2.1). Then I outline key theoretical criteria of domain-specific terminology and relate the criteria to terminology work in practice (section 2.2). Finally, I discuss the potential of combining functional lexicographical principles guiding the compilation of user-oriented specialized dictionaries with formal ontological principles representing conceptual structures of the domain-specific terminology. In particular, I propose so-called dual-entry modes to enrich term banks (section 2.3).

2.1 Introduction

The term “terminology” designates three concepts depending on the level of abstraction: Firstly, terminology is the “set of terms of a particular special subject,” (Cabré, 1999), which in practice means a “set of designations [...] belonging to one special language,” (ISO, 2000), i.e. a vocabulary. Secondly, terminology is the “guidelines used in terminographic work” (Cabré, 1999), where terminography is “part of terminology work [...] concerned with the recording and presentation of terminological data,” (ISO, 2000), i.e. a practice. Thirdly, terminology is the “principles and conceptual bases that govern the study of terms,” (Cabré, 1999), in practice we are dealing with “science studying the structure, formation, development, usage and management of terminologies [...] in various subject fields," (ISO, 2000). In other words, terminology is a vocabulary, a practice and a theory, and that confuses the literature search, because only the context will reveal which of the three meanings are meant.

My research comprises all three levels of abstraction, as I investigate experimental approaches to domain-specific terminology and term banks in a chosen domain, which extends the body of relevant literature to research on specialized language and communication emphasizing the specialized vocabulary, as well as knowledge engineering and computational linguistics emphasizing the structuring and presentation of terminological data in term banks.

My research is also guided by a functional approach to user-adaptive visualizations. The functional approach to specialized lexicography focus on the functions of dictionaries, i.e. satisfying specific needs of potential users, and terminology (and specialized lexicography) facilitates communication in specialized domains (Fuertes-Olivera, 2012). Adaptive software visualizations match the level of detail of each programming construct with users' knowledge (Loboda & Brusilovsky, 2010). It has been shown that adapting the level of explanation to users (additional explanations to novice users and specific details to expert users) result in faster comprehension and lower error rates (see e.g. Boyle & Encarnacion, 1994 and Kobsa, Koenemann & Pohl, 2001).

Dual strategy to present terminology

Term banks are used to present the meaning of terms to target users. Terms comprise a linguistic as well as a conceptual side (Roche, 2006), which renders possible a dual strategy to disseminate knowledge to target users combining the lexicographic dictionary entry mode (in textual format) with the terminological ontology entry mode (in graph format). This combinatory framework potentially enhances target users' understanding because the term bank contains complementary information on the conceptualization, i.e. the underlying concepts, relations and characteristics as well as the lexicalization, i.e. the specialized vocabulary of the discourse. However, research on interface term banks including design of entry modes has received little attention in the terminology literature, and it is necessary to include the literature with different approaches to terminology such as user-oriented lexicography and knowledge engineering.

Advances in information technology have carried lexicography into the digital age. Dictionaries are no longer constrained by paper based versions, but rather digital information tools containing rich representations (see e.g. Lew & de Schryver, 2014; Fellbaum, 2014 and Verlinde, Leroyer & Binon, 2009). The challenge of electronic dictionaries is to ensure that user needs are met and that technological features (e.g. user adaption) are also incorporated (Kwary, 2012) otherwise digital dictionaries are really no different from the paper based counterparts. Lexicography is very much devoted to general language (missing domain-specific entries), to translational and communicational needs for producing a document rather than cognitive needs of knowledge acquisition, to purely textual formats (not multimedia), and often expert users are disregarded (learners' dictionaries). Knowledge engineering and terminological ontologies have the potential to alleviate these weak points of lexicography.

Terminological ontologies represent concepts, relations and characteristics in graphs. Ontologies allow for the formal descriptions of the lexicon, which should be accessible to users (L'Homme, 2014). Ontologies possess strong advantages for humans, as consistent definitions may be inferred from the structure. But ontologies may also be used to ensure compatibility between information technologies (Madsen, 2006). There are various ways of visualizing ontologies (Katifori, Halatsis, Lepouras, Vassilakis & Giannopoulou, 2007) but lexicographers have also recognized the potential for supporting the textual lexical information of dictionaries with graphs (Polguére, 2014).

Expert target users

In theory, it is ascertained that terms should be recognised, fixed and disseminated with the help of the expert community (Cabré, 2003). The criterion rests on the assumption that only experts build knowledge and develop existing as well as new concepts, which is questionable per se, but another question also remains; why an expert is needed to validate the potential designations of concepts. In particular, should popular terms be deprecated, because they are not recommended by an expert community? The primary reason for deprecation is to avoid admitting ambiguity in the specialized communication, which may confuse target readers. However, strong motivation for term variation may exist, because the change of term variant may ensure that knowledge is conveyed to a larger group of target receivers (see chapter 3). In the taxation domain, it is often the case that the introduction of new tax types is covered by the media, where journalists choose a more popular term variant to reach their target readers, which is likely to differ from the terminology of the formal documents underlying the legislation process.

Instead of worrying about normative aspects of terms and the validation by experts, terminologists or knowledge engineers, it is important to pursue a strategy ensuring a rich representation of the term variation existing in the specialized discourse. Any member of discourse, even the flawed or less reliable sources, should ideally have a place in the term bank. Why would a proponent of unambiguous communication welcome such obvious semantic inconsistencies? Because the term bank is in fierce competition with other resources on the internet and the way to attract target users to the term bank is to meet the information overload by managing, structuring and presenting the (sometimes conflicting) discourse effectively and efficiently. A survey from the Swedish national term bank (Rikstermbanken) even reveals that users prefer access to the entire list of results instead of having the terminology center remove or merge of entries representing the same concept (Henrik Nilsson, personal communication, October 24, 2011). A reason may be that users increasingly apply habits from interaction with online search engines (Lew & de Schryver, 2014).

There is an inherent problem with attracting experts as users of the knowledge tools, which is caused by the time lag between knowledge formation and the compilation of documents for the term bank. Experts have little interest in delayed historical information about their field (Fuertes-Olivera & Nielsen, 2011). Recent research suggests rapid methods to validate draft structures (Thomsen, 2012b), which may provide the necessary component for attracting experts as users, but equally important attracting young generations not familiar with paper based reference works, who prefer other online sources (e.g. Google).

Term-bank user interface

The behaviour of dictionary users is likely to resemble the behaviour of term-bank users, especially in the case of rich, electronic and multimedia dictionaries. In general, lexicographic researchers emphasize dictionaries aimed at users with needs to translate general language, despite the potential to cover other user needs such as specialized communication, cognition and learning (Caruso, 2011), and many domain-specific terms are not included in the dictionaries (see the 'n.a.' entries in column 3 of table I.2 in appendix I).

Research on dictionary interfaces falls in three broad groups: recording or logging actual user behaviour in a natural setting (Müller-Spitzer, Wolfgang & Koplenig, 2015), recording user behaviour in a controlled experiment (e.g. by using eye-tracking technology), or rating of interfaces detached from any recording of behaviour (Lew & de Schryver, 2014). So far user adaption and multimedia access have received poor ratings by users, presumably reflecting users who are unfamiliar with modern features of dictionaries (Lew & de Schryver, 2014). Therefore, we need a controlled experiment with manipulated questions ensuring that participants search across the entire proposed dual-entry mode (see chapter 7).

Scale of specificity

We may assume that the linguistic units of a specialized text consist of core domain-specific terms (e.g. “fiscal taxes”, see figure 2.1), non-core terms not specific to any particular domain or weakly specific to a range of domains (e.g. “consumption”, see figure 2.1) and a residual containing general-language words or phrases (e.g. “a”, “are”, or “the”, see figure 2.1). In this framework, the domain-specific terms are the most specific and distinct vocabulary belonging uniquely to the domain and possess the highest level of specificity. An increasing level of specificity allows for the presentation of semantic relations in so-called wordnets (see e.g. DanNet, 2015) and possibly also in formal terminological ontologies visualized in graph format (see section 2.3.1).

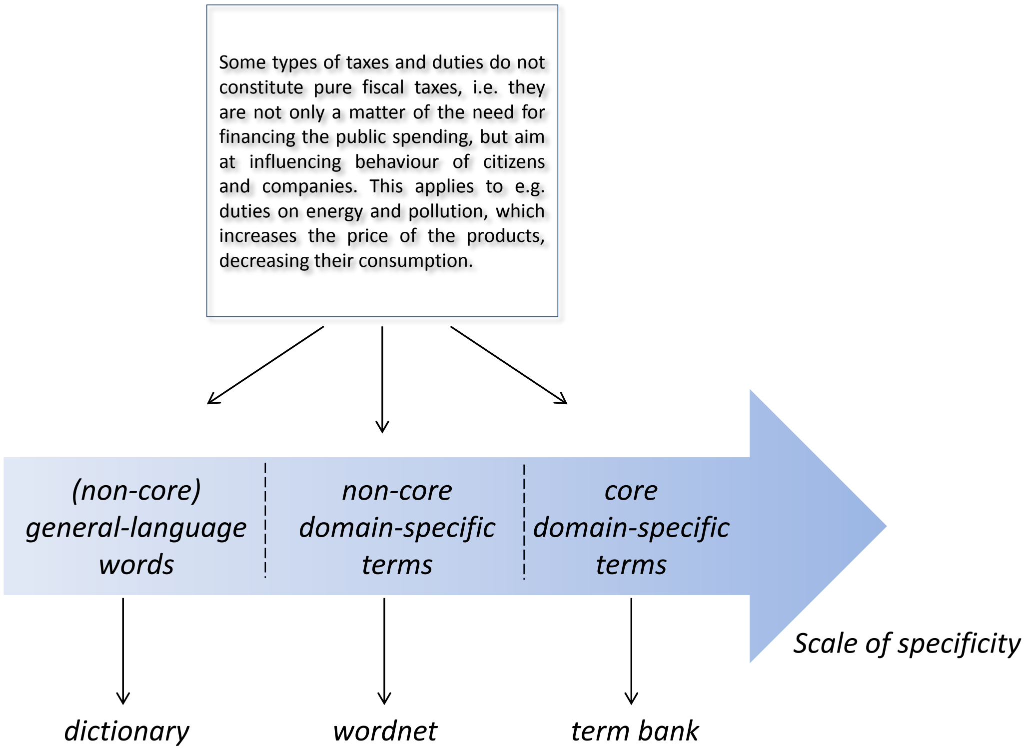


Figure 2.1: The scale of specificity of linguistic units. In a translated specialized text (see figure C.3 in appendix C for the entire Danish source text), we identify three groups of linguistic units: Core domain-specific terms constitute the highest level of specificity and are typically presented in term banks. Non-core domain-specific terms exhibit a lower level of specificity and may be presented in so-called wordnets illustrating semantic relations. The remaining general-language words are non-specific (and non-core) words are typically presented in dictionaries with semantic and linguistic information.

It is difficult to separate general and specialized texts and therefore to distinguish a specialized discourse from a general discourse, see e.g. Da Cunha, Cabré, San Juan, Sierra, Torres-Moreno & Vivaldi (2011) for a review. The existence and ranking of the three groups of specificity in figure 2.1 is beyond doubt. However, the exact distinction between what we may label the non-core domain-specific terminology from the core domain-specific terminology may prove problematic, which is illustrated by dashed lines. The same goes for the distinction between the non-core terminology and some general-language words (especially nouns), because general language may have specialized usage in a particular domain (Nkwenti-Azeh, 1994).

Despite the shared features of term banks and dictionaries or semantic networks (wordnets), I infer from my literature review that no other knowledge tool allows for a rich dual representation of the domain-specific terminology in the user interface, i.e. lexical and conceptual information visualized by text and graphs with potential for actually meeting diverse user needs (cognition, communication and translation). In particular, term banks allow us to for target expert users exhibiting high level of expertise.

2.2 Domain-specific terminology

In this section, the fundamental theoretical criteria of terminology are presented and discussed in relation to the domain-specific terminology using taxation as case study.

Terms are the object of study of terminology scholars. As noted by Cabré (2003), it is indispensable that terminology as a distinct and independent science is able to identify and account for the specificity of the terminological units. However, the theoretical principles on which terminology science rest appear vague, because terminology originates from Wüster's standardization work in the domain of engineering (Wüster, 1968), which necessarily emphasizes a strong practical side (Wüster, 1974).

Terms are assumed to cover three theoretical criteria (see e.g. Cabré, 2003). A communicative criterion because terms further a specialized discourse (see section 2.2.1), a linguistic criterion because terms constitute lexical units (see section 2.2.2) and a cognitive criterion because terms designates underlying concepts (see section 2.2.3). According to Cabré (2003), terms may be approached from the view of any of the theoretical criteria as no particular order is necessary. In my view, systematic terminology work is closely related to the specialized discourse i.e. identifying the relevant texts, retrieving the appropriate lexical units and the characteristics of the underlying concepts As the criteria are outlined, I relate them to the Danish taxation terminology in a case study, where the discourse is delimited to written documents (i.e. texts) comprising what we may label the "three phases" of revenue collection beginning with the financial appropriation (i.e. the motivation for imposing a tax) and subsequently legal ratification of direct and indirect taxes (i.e. the establishment of authority necessary for collecting the tax) to revenue statistics (i.e. assessment of the actual taxes paid). The result of the domain-specific terminology work is shown in appendices I and II.

2.2.1 Terms furthering specialized discourse

Terms can be identified as units of communication (Cabré, 2003), which implies that terms occur in authentic specialized discourse.

Specialized discourse may in principle be multi-modal, however, specialized discourses exhibit a level of formality and accuracy, which produces a preference for certain written text types, which are particularly suited for a systematic presentation of the specialized knowledge. Two features are distinct for the specialized texts: a lexical feature reflecting that the domain-specific terms are exclusive to the topic and having a narrow meaning in the particular context and a textual feature reflecting a more concise and systematic expression than general texts achieved by the use of appropriate grammatical devices (Cabré, 2003). We may suggest that terms are crucial to ensuring unambiguous specialized discourse. In the following, I describe the chosen distinctive features of the taxation discourse, which is formal, legal and financial.

Formal taxation discourse features

Seen from a terminological point of view, a discourse analysis of the taxation domain reveals that the modalities covered by the revenue collection discourse from appropriation over budget to national accounting constitute written modes including schematic or graphic representations of the size, structure or source of the revenue. These illustrations provide us in part with systematic grouping and hierarchy applicable in the subsequent conceptual structuring of taxation concepts. However, it should be noted that neither in existing term banks, nor in existing tax reference works are graphs depicting conceptual structures (including formal feature specifications) available (see table 2.2).

Moreover, the taxation discourse comprises written documents of high formal status (motivating, collecting or assessing taxes). The taxation discourse is often very difficult to verbalize, comprising long, multiword, infrequent, inpredictable terms e.g. “third party liability insurance fiscal act” (*motoransvarsforsikrings-afgiftsloven*). The taxation discourse is highly abstract and difficult to embody into images, and illustrations usually show tax payments or process diagrams underlying the technologies for collecting tax revenues. Not the terminology per se.

The limitations to the modes representing the taxation discourse are due to the nature of taxation being a highly abstract (or fictitious) construct, which is not materialized into objects available to our senses (or cameras) and therefore not able to photograph or video record.

Legal taxation discourse

The taxation discourse constitutes legal terminology in need of exact precision and reference (see e.g. Sandrini, 1999). The legal discourse is Danish, as the language of the courts is Danish, and that shapes the tax terms accordingly (*Retsplejeloven § 149*, Retsinformation, 2015a). The taxation discourse is to a large extent to be found in Danish as the source language (SL), referring to Danish laws, leaving very little room for interpretation or translation freedom to the target users of the domain. Strictly speaking, taxation is a matter of taxpayer liability, which is defined by the legal text.

Legal terminology carries the risk of expiration due to the abolishment or amendment of the law defining or administrating a particular tax. Expired texts should be clearly marked, which leaves responsible authorities with a serious challenge in ensuring processes that will effectively, and perhaps automatically, monitor for expiration and will update relevant information (available in term banks) accordingly. However, from a confined terminological point of view, it does not necessarily constitute a problem that a term is abolished, on the contrary, terminated information may contain valuable knowledge necessary for concept clarification, even showing the holes which follow logically from the formal terminological ontologies e.g. the abolished “middle-bracket tax” (*mellemskat*) fitting into the structure between bottom and top-bracket taxes. It can, however, not be stressed enough that a term bank providing unmarked terminated information will undermine its own reliability, which poses a serious challenge to term banks of the digital era, where knowledge resources are fiercely competing for users.

Financial taxation discourse

A particular distinctive feature of the taxation discourse is the financial aspect, which is shown by the many figures representing revenue amounts, tax rates, tax base units in volume or weight. This feature may prove useful as an additional informative attribute in evaluating synonymy or equivalence across languages if e.g. the statistical reference (revenue or rate) is approximately identical (see the revenue column of table I.1 in appendix I). However, there may be good reasons why numbers, which we in theory expect to be identical, do not match, which complicates concept clarification that is based on quantitative measures. First of all, mistakes often happen especially in reproducing numbers to other currencies (e.g. Danish Kroner to euro) or converting actual numbers to other units (from million or billion). Secondly, time lags in publications, especially between budget and balance sheets, might produce discrepancies among what is expected to be realized one future point in time, and what is actually being reported years later for that particular point in time.

To sum up, the taxation discourse is shaped by legal and financial aspects and that adds challenges of liability expiration and monetary discrepancy. Moreover, the taxation discourse is highly formal even though new less formal perspectives are important to the compliance efforts to make taxpayers' understanding and willingness to pay increase. Not only do expert communities produce taxation discourse and more popular discourse with non-experts as target readers proliferate to support the digital self-filing of taxpayers.

In practice, I select my discourse form the following authorities: Firstly, the legal basis in which the field of responsibility is the Danish Ministry of Taxation (*Skatteministeriet)* (Retsinformation, 2015b), but also departmental notices and instructions published by SKAT (SKAT, 2014). Secondly, the part of the national budget (i.e. section 38) estimating the revenue of the central government taxes published by the Danish Ministry of Finance (*Finansministeriet*) (Ministry of

Finance, 2014). Thirdly, the national accounts of the tax revenue which is published by Statistics Denmark (*Danmarks Statistik*) in a bilingual (Danish and English) report (Statistics Denmark, 2014), and fourthly, the annual publication Taxation Trends in the European Union which is based on national accounts and published by the statistical office of the European Union (EuroStat, 2014). To some extent, the revenue statistics published by the OECD is included (OECD, 2014).

Danish taxation terminology, in particular, and domain-specific terminology, in general, is prone to non-equivalence. Consequently, it often becomes impossible to find the specialized texts in target language (TL), which contains equivalent terms as they are non-existing. In those cases, a more effective translation strategy is to translate an explanation of the term, i.e. a definition (see table 3.2). It is argued that the selection of equivalents should be of secondary consideration compared to the definition (Nkwenti-Azeh, 1994).

Definitions are crucial

The taxation discourse may provide us with definitions, context examples and equivalent terms. However, definitions are of primary concern to terminology work and key to conveying meaning of terms and specialized knowledge to target users. Apart from the term itself, a central information category in term banks is the definition (Thomsen, 2012a). In the compilation of dictionaries, emphasis is also put on the definition as the prototypical carrier of meaning (Lew, 2010). The lexical and conceptual criteria underlying terms provide for two kinds of term definitions:

First, the verbal explanations, which can be extracted from the discourse by means of context examples, which paraphrases, exemplifies or elaborates the term. For instance, “energy tax” is elaborated by “The largest subgroup of the excise duties is energy taxes, which constituted 44 per cent of total excise duties in 2011” (*den største gruppe blandt punktafgifterne er energiskatterne, der i 2011 tegnede sig for 44 pct. af samtlige punktafgifter*) (see figure D.3 in appendix D). Indeed, these “technical” definitions may contain linguistic relations (i.e. hypernymy, hyponymy, synonymy and meronymy) to other terms (Roche, 2006; Roche, Calberg-Challot, Damas & Rouard, 2009). However, the quality of the technical definitions is highly dependent on the criteria underlying the selection of the examples (Lew, 2010). In some cases, the distinction between terms and their definitions may be blurred because terms may be long and complex and almost serve as definitions, and conversely some definitions are so short, they could almost be thought of as terms (ISO, 2009).

Second, the logical specifications of the underlying concepts, relations and characteristics, which may also be extracted from the discourse or preferably inferred from the conceptual structures (Roche, Calberg-Challot, Damas & Rouard, 2009). For instance, we can define "energy tax" (*energiskat*) as a (type of) excise duty (hypernym or super-ordinate) imposed on goods and services with the purpose of limiting the environmentally damaging energy consumption (delimiting characteristics) (see figure D.3 in appendix D). These "conceptual" definitions are consistent because they comply with the conceptual organization of the term bank. (We return to the conceptual definitions in section 2.2.3.)

2.2.2 Terms constituting lexical units

Terms can be identified as units of (specialized) language as they contain a linguistic side (Cabré, 2003). However, the lexical unit of terms has a preference for certain word classes as they often occur as nouns, verbs, adjectives or adverbs (Cabré, 2003). The identification of terms in a specialized document is challenged by the fact that terms formally coincides with units of the general language or general discourse (Cabré, 2003), which supports the critique language science raises with the question of a clear separation (and definition) of specialized language from general language. In particular, the linguistic characteristics of terms (e.g. phonological, morphological and syntactic characteristics) are not necessary different from other lexical units (Cabré, 2003) (see figure 2.1, where terms are not odd). L'Homme (2003) confirms that it is not possible from a purely linguistic point of view to distinguish terms, since the "specialized status" is determined by the relation to a given subject field, i.e. terms contain certain specialized content. It could be argued that a lexical unit is general by default, but acquires the specialized or terminological meaning when it is activated by the pragmatic characteristics of the discourse. Any lexical unit would thus have the potential of being a term (Cabré, 2003).

Odd lexical units

It is not necessarily required that terms follow the morphological (word) or syntactical (phrase) rules of a particular language (the structure of a given language's morphemes and other linguistic units, such as root words, affixes, parts of speech, etc.). Indeed, terms may take forms that are not fully complying with the morphological (or ortographic) ruling, e.g. "act on investor funds" (*lov om investorfonds*) ⟐, "Income and Wealth Tax to the Central Government" (*Indkomst- og Formueskat til Staten*) ⟐, "pure fiscal tax" (*ren fiskal skat*). It is also often the case that acronyms are used terms. Even inside a small domain, homographic acronyms exist, e.g. CFC is used both in connection with tax on "controlled-foreign-companies" as well as "duty on chlorofluorocarbon."

Formal approaches to term extraction will identify terms (including intra-term relations and inter-term relations) that follow certain morphological and syntactic patterns e.g. "-tax" (*-skat*), but this will not capture all the terms present in the discourse, especially the terms that appear embedded with a low distance to natural language e.g. "persons liable to tax in this country" (*personer, der er skattepligtige her til landet*). Conversely, terms that contain "odd" components from other languages e.g. "A-tax" (*A-skat*), i.e. tax deducted from income at source, are not necessarily captured either. When terms stand out, they are easy to identify as they constitute odd lexical units different from words or phrases (see table 3.2).

Conceptual entries

In a dictionary, article entries are devoted to words (lemmata), no matter whether they constitute lexical units, or merely morphological units. If I have identified term candidates that are members of a specialized discourse (not necessarily carrying specialized meaning) and that (more or less) comply with typical morphological and syntactical ruling of a particular language, we are not necessarily opening a term-bank entry. To the lexicographer, this would be sufficient information to open a dictionary entry, but to the terminologist the motivations behind the lexicalization are crucial as the pragmatic term variants may limit the allowed forms. We may e.g. ask ourselves how we should divide "taxes and duties" (*skatter og afgifter*). Is it a tax or several taxes (in plural), or is it necessary to distinguish between "tax and duty", or perhaps even "tax, VAT and duties," as the object of the taxation domain? To determine the question of the meaning, we need to inquire into the underlying concepts. However, lexicalization (or term variation) is not necessarily conceptually motivated as we shall see in chapter 3.

Multilingual terminology

No thesaurus (monolingual) or specialized dictionary (bi- or multi-lingual) covers the Danish taxation system in detail, despite the fact that taxation is a matter relevant to minimum the 88% taxable members of the Danish population (Kleven, Knudsen, Kreiner, Pedersen & Saez, 2011). The culture-specific taxation domain makes it very difficult to translate terms into other languages (see e.g. Sandrini, 1999), where different legal institutions are working. In some cases, we are aiming for a translation into British English, and then we must look into Britain's national taxation system. In other cases, we want to communicate at supranational levels, either regionally to the EU, or globally to the OECD. The annual publication by EuroStat "Taxation Trends in the European Union" discusses taxation structures. This publication does not fit the Danish counterpart published by Statistics Denmark, and compared to the OECD counterpart "Revenue Statistics" introduces even more equivalence problems.

2.2.3 Terms designating concepts

Terms can be identified as units of (specialized) knowledge as they contain a cognitive or conceptual side (Cabré, 2003). The final criterion enabling us to distinguish terms from other lexical units of the specialized texts is that terms contain specialized content i.e. occupy a precise place (node) in a conceptual structure of a subject field (Cabré, 2003). The specialized discourse presents an organised structure of knowledge, where concepts, relations and characteristics are expressed by means of linguistic units. Concepts emphasize an underlying system, which does not necessarily appear from the pure linguistic components of a term, because terms and the underlying concepts belong to different semiotic systems (Roche, Calberg-Challot, Damas & Rouard, 2009). In practice, we are able to construct the terminological ontologies by means of linguistic units (textual cues) extracted by interpreting the source material. In particular, we may model domain-specific concepts in so-called terminological ontologies by formal feature specifications of attribute-value pairs (Madsen, Thomsen & Vikner, 2004).

Concept clarification

The concept clarification process of the term “personal income tax” (*personskat*) is shown in table 2.1. A manual term extraction is resulting in a list of lexical units from the relevant discourse (see column 4) comprising multiword terms in various patterns (see column 3), from which the concepts, relations and characteristics are inferred (see column 2). The table illustrates an example of domain-specific terminology complying with the three fundamental criteria of terminology:

Table 2.1: Concept clarification (see table I.1 in appendix I for the full version). Column numbers refer to the columns in table I.1.

Discourse (source)	Lexical (term)	Concept (characteristics)	Legal (act)	Financial (amount)
(4)	(3)	(2)	(5)	(6)
SKM	*indkomstskat for personer*	Taxpayer: Liable persons	LBK 143	N.a.
FFL	*personskatter*			227.5
DST	*personlige indkomstskat*			232.9

Discourse (source), lexical units, and underlying conceptual content. Moreover, the distinctive features of the taxation discourse (legal aspects with reference to specific laws (column 5) and revenue size (column 6)) are shown.

Terminological ontology

The conceptual side of terms allows us to model terminological ontologies from the formal feature specifications (characteristics) (DanTermBank, 2015b). Figure 2.2 shows an example (see figure II.10 in appendix II). It is possible to infer definitions from the terminological ontology by means of relations and characteristics. The terminological ontology in appendix II (see figure II.1 for the entire terminological ontology of the Danish tax system and figures II.2-II.19 for extracts of each part of the entire ontology) has been created using the concept modelling module i-Model of the terminology and knowledge management system i-Term ® developed by the local terminology centre at Copenhagen Business School, which allows the user to construct terminological ontologies from entered domain-specific information (Madsen, Thomsen, Halskov & Lassen, 2010).

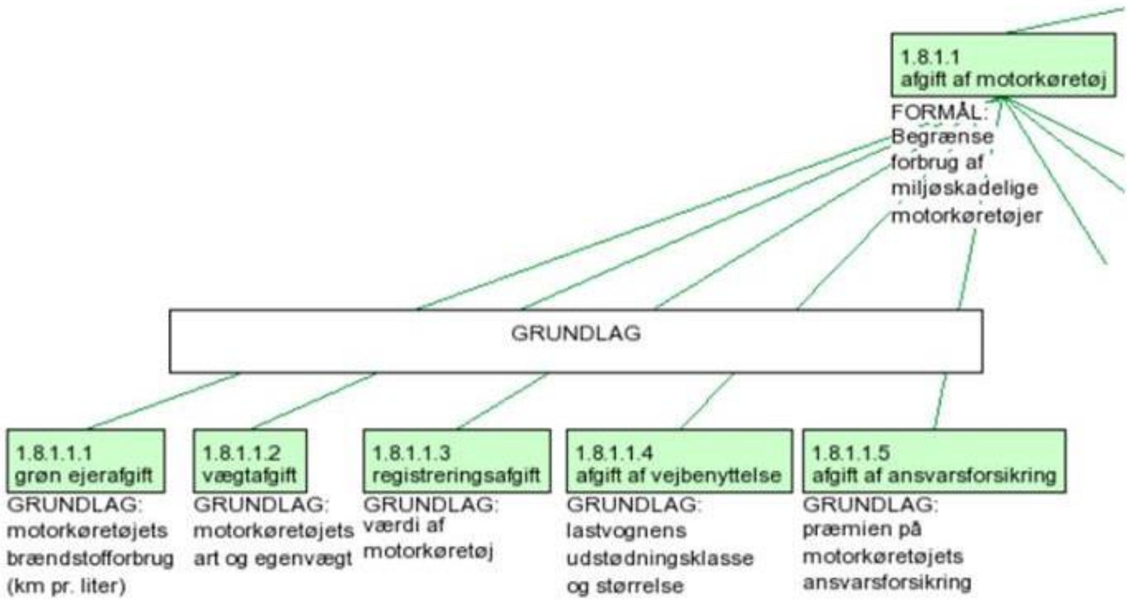


Figure 2.2: Domain-specific terminological ontology. Extract of ontology of taxation terminology, where the characteristics of the concepts are presented as feature specifications in the form of attribute value pairs, e.g. 1.8.1.1.2 “weight tax” with TAX BASE: the type and weight of the vehicle. Subdivision criteria are introduced on the basis of the feature specifications (white boxes with text in capital letters) which illustrate that the five shown coordinate concepts 1.8.1.1.1- 1.8.1.1.5 differ with respect to tax base (*grundlag*).

In practice, this means that e.g. “weight tax” (*vægtafgift)* inherents the features of the superordinate indirect tax (mother) “motor vehicles tax” (*afgift af motorkøretøj*), and is distinguished from sideordinates (sisters) by means of “tax base: the type and weight of the vehicle” (*grundlag: motorkøretøjets art og egenvægt*). The modeling language is formal, but the specification of features is inferred from specialized language, which is used in the terminological ontology.

The concept clarification underlying the terminological ontologies may reveal non-correspondence in the structures, which could be a case of non-lexicalization (Mourier & Vesterli, 2009) of missing concepts appears. For instance “duty on stimulants” (*afgift af nydelsesmidler*) was missing in structure (see Figure II.18 in appendix II). Moreover, aligning several terminological ontologies may also reveal non-correspondence or different structuring principles, which reflects non-equivalence between source and target languages.

Conceptual graphs

Terminological ontologies allow for a graphical visualization, which extends the dissemination media of term banks beyond the textual (and hypertextual) specialized dictionary format. In the lexicographic (semasiologic) approach, the content of each entry subsumes all (polysemic) meanings of a particular entry shaping the structure of each entry (micro-structure), which is opposite to the terminological approach, where the focal point is the conceptual side designated by (synonymous) terms. Moreover, the onomasiologic approach renders possible a conceptual systematic structure of the term bank (macro-structure) in addition to the pure alphabetical one, and conceptual relations will benefit the internal structures (medio-structures) producing strong cross-referencing from which consistent definitions can be inferred. Finally, the discourse underlying terms allows for the compilation of highly relevant context examples from the actual specialized discourse.

Inferring definitions

The underlying concepts, relations and characteristics of terms render possible the inference of (intensional) definitions from the conceptual structures in the terminological ontology. But other definition types exist, as mentioned technical definitions can be extracted directly from source material. Despite the fact that these definitions are not consistent as they are not inferred from the terminological ontology, they may prove very important carriers of meaning to target users.

According to ISO (2000), the definition is the representation of a concept by a descriptive statement which serves to differentiate it from related concepts. In the taxation domain, the intensional definition type is the most useful, because it describes the intension of a concept by stating the super-ordinate concept and the delimiting characteristics. In the formal terminological ontology, definitions are directly inferred from the structures, and following this guideline, we may e.g. define “personal income tax” (*personskat*) as a (type of) direct tax (super-ordinate) imposed on liable persons (characteristics).

2.3 Term banks

In this section, I propose the combination of functional lexicography, which guides the compilation of user-oriented specialized dictionaries, with formal ontological principles, which visualize conceptual structures underlying the domain-specific terminology. Moreover, I introduce a dual-entry mode assigning the term-oriented textual entry to the relevant concept-oriented graphical entry, which has the potential of enriching term banks to the benefit of target users. This dual strategy is rare in the terminology field.

Term-bank evolution

In their beginning (mid 1960s to early 1970s), term banks were closely linked to the field of translation (Nkwenti-Azeh, 1994). However, the technological development following and the transformation of dictionaries into electronic formats has blurred the distinction between terminology and lexicography as both disciplines are aimed at compiling linguistic data. However, term banks have the potential of being more than a specialized dictionary, if we apply the formal and consistent ontological principles. Term banks may not only facilitate the acquisition of terminology and knowledge of (human) target users, they may also contain data that are readable to a computer (Roche, 2006).

Nkwenti-Azeh (1994) describes the three generations of term banks starting with the first generation of term-oriented printed technical (or specialized) dictionary, to the second generation of concept-oriented and the third generation of knowledge-oriented terminology tools. Term banks are displaying computerized terminology to target users, who face a serious time constraint if they were to conduct systematic terminology work presented in the term banks. Up to 60% of total time used for specialized translation is devoted to terminology work (Nkwenti-Azeh, 1994). The conceptual side of terms is what allows the terminologist to separate terms from words (Nkwenti-Azeh, 1994) and renders possible a systematic order of term-bank entries organised by the concepts (Thomsen, 2012a). However, a review shows that many online dictionaries are unsystematic and lack consistent organization (Caruso, 2011).

Iterative approaches

Terms rest on the three fundamental criteria: Terms are members of a specialized discourse, constitute lexical units and designate an underlying concept. The chosen order of applying the fundamental criteria in the case study resembles a word-oriented semasiologic approach (Thomsen, 2012a). Alternatively, we could have chosen a concept-oriented onomasiologic approach (Thomsen, 2012a) by starting with a group of taxation experts who identified and structured the concepts of their domain, and then we could have identified the designations and compiled a corpus from the documents containing those designations. It should be noted that the former semasiologic approach focuses on terms and allows for polysemy (one term may designate different concepts, even within one domain), while the latter onomasiologic approach focuses on concepts and allows for synonymy (different terms designate the same concept).

According to Cabré (2003), this does not constitute a problem to the terminology field, as terms can be accessed through different “doors,” which means that terminologists are free to choose from the two approaches as long as they keep all three fundamental criteria in mind. It is noted by Fernández-Silva, Freixa & Cabré (2011) that the most frequent way of approaching concepts and the underlying conceptual structures is by means of their linguistic representations in specialized texts. In my view, the two approaches need not be mutually exclusive, or rather one should aim for an iterative practice to account for the dynamic nature of domains as well as gap-filling practices sorting out the concepts logically missing in the structure of complex domains.

Dual-entry mode design

Mahalingam & Huhns (1998) claims that “one of the attractive features of graphical representations is that they are easier to understand than their textual counterparts.” Indeed Frappier, Meynell & Brown (2013) emphasize that knowledge can be represented by means of e.g. diagrams by applying graph theory.

It is my aim to develop a dual-entry mode combining the term-oriented (to be visualized in text) with the concept-oriented (to be visualized as graphs) to convey domain-specific terminology and knowledge to target users. A review of typical examples of existing term banks, glossaries and ontologies (see table 2.2) reveals that the duality is rare. It is possible to identify resources applying both visualization modes, but there is a preference for the textual mode. The primary information mode available to users of conventional term banks (row 1-3 in Table 2.2) and glossaries specific to the chosen domain (row 4-7 in table 2.2) is merely applying a textual schematic information mode, while the Danish (non-domain specific) Wordnet (row 8 in table 2.2) also applies the graphical information mode depicting conceptual structures. Finally, examples of advanced knowledge technologies such as the monolingual (Danish) health care terminology (row 9 in table 2.2) and the multilingual “WIPO” terminology covering many domains (10) provide searchable access into conceptual structures displayed in a graphical mode with navigational functionalities, in addition to the pure textual information.

Table 2.2: Terminology visualizations in selected terminological resources. Typical examples of term banks and glossaries are evaluated with respect to the textual and graphical displays. First, term banks in Scandinavia in the form of a national term bank in Sweden (1), a term wiki in Norway (2) and the broad term bank of University of Vaasa in Finland (3) are relying on textual displays. Second, glossaries covering tax policy in the US (4), statistics of EuroStat (5), taxation terms of the OECD (6) and financial terms of the World Bank (7) are also relying on textual displays. Third, the Danish wordnet (8), the Danish health care term bank (9) and the WIPO's multilingual terminology portal (10) relies on both the textual and graphical displays.

	Origin	**Content**	**Textual display**	**Graphical display**
1	Sweden	National term bank (rikstermbanken.se)	Yes	No
2	Norway	Terminologi og fagspråk (termwiki.sprakradet.no)	Yes	No
3	Finland	University glossaries (uva.fi)	Yes	No
4	TPC	Tax glossary (taxpolicycenter.org)	Yes	No
5	EuroStat	Statistic glossary (eurostat.ec.europa.eu)	Yes	No
6	OECD	Tax terms (oecd.org)	Yes	No
7	World Bank	Glossary (worldbank.org)	Yes	No
8	DanNet	Lexical ontology (wordnet.dk)	Yes	Yes
9	SSI	Health care terminology (begrebsbasen.sst.dk)	Yes	Yes
10	WIPO	Multilingual terminology (wipo.int)	Yes	Yes

The duality of the entry modes of a term-bank user interface is novel, and the combinatory method will be outlined in section 2.3.3. It should be noted that the experiments (see chapter 7) will reveal, whether the prototypes are effectively and efficiently conveying knowledge to target users and should be maintained in the subsequent user interface design process, or whether the proposed dual-entry mode should be rejected.

2.3.1 Terminological ontologies

In this section, I present the practice of knowledge engineering in proposing the graphical part of the dual-entry mode. The conceptual side of terms renders possible the building of terminological ontologies. It is possible to apply "formal ontological principles in the current practice of knowledge engineering" (Guarino, 1995, p. 627).

Roche (2006, p. 1034) ascertains the promise of ontologies to be the "way of capturing a shared and common understanding of a domain that can be understood and used as well by humans as programs." However, the introduction of ontology into terminology work and knowledge engineering is not easy, as the definitions of ontology vary. Roche (2006, p. 1034) suggests that "an ontology is a shared description of concepts and relationships of a domain expressed in a computer readable language [...] it will include a vocabulary of terms and some specification of their meaning (i.e. definitions)."

Formal ontologies

Ontology is borrowed from philosophy, where it is defined as "a systematic explanation of being" (Corcho, Fernández-López & Gómez-Pérez, 2003) or "a systematic account of existence" (Gruber, 1995; Jean, Pierra & Ait-Ameur, 2007). Moreover a distinction is necessary between "ontology" as the branch of philosophy and "an ontology" as a classification of categories (Corcho, Fernández-López & Gómez-Pérez, 2006). In the ontology community, Gruber (1993) defined an ontology as the explicit specification of a conceptualization. Borst (1997) replaced the explicit with a "formal" specification and emphasized that the conceptualization should be "shared" by computers and humans to facilitate knowledge sharing (Flahive, Taniar, Rahayu & Apduhan, 2011). Jean, Pierra & Ait-Ameur (2007) distinguishes ontologies from conceptual models, which "respect the formal criterion" and is "based on a rigorously formalized logical theory."

Generally speaking, an ontology is the "knowledge representation medium" (Brewster & O'Hara, 2007) and captures the existing "things" in the application domain by "encod[ing] human knowledge and reasoning by symbols that can be processed by a computer to obtain intelligent behavior," Chein, Mugnier & Croitoru (2013, p. 249). In practice, Roche (2006, p. 1037) uses a logic-oriented formal language based on the so-called specific-difference theory, for which "a conceptualization is a system of concepts organized according to their differences (a concept is defined from a previously existing one by adding a specific difference)" for the definition of a conceptualization. Or put differently by Noy (2004), "an ontology is some formal description of a domain of discourse, intended for sharing among different applications, and expressed in a language that can be used for reasoning." The aim of ontologies may be stated as capturing "consensual knowledge in a generic and formal way, and that they may be reused and shared across applications (software) and by groups of people" (Corcho, Fernández-López & Gómez-Pérez, 2003, p. 44). Finally, Brewster & O'Hara (2007, p. 564) emphasize that there is a "whole range of functions, assumptions and aspirations encoded in a given type or instance of an ontology."

Domain-specific terminological ontologies

Jean, Pierra & Ait-Ameur (2007, p. 240) introduce domain ontology as a "domain conceptualization" defined as "a formal and consensual dictionary of categories and properties of entities of a domain and the relationships that hold among them." It is widely agreed that there is a need to "create an explicit specification of how knowledge in a domain is conceptualized," (Flahive, Taniar, Rahayu & Apduhan, 2011, p. 618). Domain-specific ontologies are terminological ontology or concept systems that specify the meaning of domain-specific concepts in the form of characteristics and relations to other concepts. According to Madsen (2006, p. 2) "[t]erminological ontologies are based on an analysis of characteristics that the terminologist obtains from texts as well as from discussions with subject experts."

The structuring advantage from terminological ontologies should be seen in the light of the absent feature of the specialized dictionaries, where the so-called medio-structure between related entries is inadequate. Regarding the medio-structure, de Shryver (2003, p. 180) defines this as “known as the system of cross-referencing, is used to connect different components of a dictionary… “ and underlines that it “includes both dictionary-internal cross-references (e.g. central-section data with front and back matter material, articles with related articles, article-internal sections with one another, etc.), and dictionary-external cross-references (such as links with corpora, other reference works, the Internet, etc.”

Example of terminological ontology

I construct the terminological ontology from the chosen revenue collection discourse following the ISO-standard on terminology work (ISO, 2000). A conceptual analysis based on the technical definitions in the texts results in the extraction of 95 terms (see Table I.2 in appendix I). Firstly, the underlying concepts are structured by means of subdivision criteria “collection form” (*opkrævningsform*), see figure II.2, or “revenue receiver” (*modtager af provenu*), see figure II.3.

In the direct taxation case, the secondary subdivision criteria are “taxpayer” (*skatteyder*), see figure II.4, or “tax base” (*skattegrundlag*), see figures II.5 and II.6. In the indirect taxation case, subdivision criteria constitute also “tax base” (*skattegrundlag*), despite that tax burdens can be shifted across taxpayers, see figures II.7 and II.8. Then “purpose” (*formål*), see figure II.9, followed by different forms of content, feature, resource or product are the structuring principle, see figures II.10-II.19. The type relation is the chosen relation throughout the terminological ontology, but the revenue collection perspective might also support the part-whole relations, instead depicting the Danish taxation structure. Moreover, temporal conceptual structures could have been added to show the calculations in the assessment of the personal income taxation. It follows from the terminological ontology that the indirect taxation is much more comprehensive in both depth (more levels) and breadth (more types of taxes).

2.3.2 Specialized dictionaries

In this section, I discuss the term-oriented electronic specialized dictionaries, which may coincide with term banks in the presentation of meaning, and therefore I should consider the practice of lexicography in proposing the textual part of the dual-entry mode.

The technological development, which has transformed printed resources into electronically accessible tools, is blurring the distinction between term banks and specialized dictionaries. However, L'Homme (2003) emphasizes that specialized dictionaries cover only a specific field of knowledge, while the computerized terminological bases are much larger and potentially cover a variety of domains. We may confirm L'Homme (2003), as only few taxation terms are included in existing specialized dictionaries (Gyldendal, 2015b) with no entries for e.g. “duty on nitrogen” (*afgift af kvælstof*) or “duty on CFC” (*CFC-afgift*) (see entries marked “n.a.” in table I.2 in appendix I).

Tarp (2000, p. 196) discusses lexicographic functions to capture the users' needs, where the lexicographic function is defined as the potential of the dictionary to cover the complex of needs that arise in the user in a particular user situation. In particular, Verlinde, Leroyer & Binon (2009, p. 1) proposes a user-adapted data-access approach to ensure that functional information tools capture the information needs of their users in different situations i.e. support users with specific communication problems (i.e. reading, writing or translating) or knowledge problems (acquiring new knowledge or verifying existing knowledge, learning a language or a subject field).

De Shryver (2003, p. 178) states that it is “one thing to be able to store ever more data, but another thing entirely to present just the data users want in response to a particular look-up.” In particular, de Schryver (2003) outlines a list of unfulfilled “dreams” of applying multimedia including “various kinds of graphs” (De Schryver, 2003, table 10). Indeed, the textual preference for the description of meaning is linked to the era of printed dictionaries, while the digital format supports the use of multimedia. In other words, functional lexicography is user-oriented.

2.3.3 Combining lexicography and terminology

In this section I propose a framework for combining the user-oriented lexicography and the concept-oriented formal terminology into a dual-entry mode, where the former is guiding the textual part (see section 2.3.2), and the latter is guiding the graphical part (see section 2.3.1).

Madsen, Thomsen, Halskov & Lassen (2010) argues for a convergence between lexicography and terminology, which allows for inserting (concept-oriented) encyclopedic or lexical-semantic knowledge into the (word-oriented) user interface. According to Grabar, Hamon & Bodenreider (2012) ontologies organize knowledge, while specialized dictionaries describe less organized terms. It is likely that dictionaries maintain a preference for the presentation of linguistic information simply because the content does not always allow for strict organization (see figure 2.1). Fernández-Silva, Freixa & Cabré (2011) ascertains that terms are often formed to reflect the most relevant characteristics of the underlying concept, and that enables the formal in terminological ontologies.

“With the use of databases, however, the possibilities for presentation do not depend on the structure of the data collection, and thus it is possible to present data from a term base with a concept-oriented structure in a word-oriented user interface. Compared to the restrictions inherent in printed publications, modern technology offers unlimited opportunities with respect to volume.

Terminology-ontology complementation

Despite the distinct goals of ontologies and terminologies, they may be intended to complement each other. Roche, Calberg-Challot, Damas & Rouard (2009, p. 322) emphasize that “[a]n [explicit] ontology may take a variety of forms, but necessarily it will include a vocabulary of terms and some specification of their meaning (i.e. definitions)”. Nevertheless, we have to bear in mind that an ontology, defined as a “specification of a conceptualization”, is primarily ‘a description (like a formal specification of a program) of the concepts and relationships that can exist’ (Gruber et al. 1993). Therefore, an ontology is not a terminology,” (Roche, Calberg-Challot, Damas & Rouard, 2009, p. 322).

I should instead approach terminologies and ontologies as part of the same continuum going from the vague representation format to the consistent and formal format. Grabar, Hamon & Bodenreider (2012, p. 375) state that “a terminology is usually defined as a set of terms, which represent the system of concepts for an area or for an application. Terms are linguistic entities,” while “[a]n ontology also describes a system of concepts and its associated properties for a specific area.” and “ontologies are built upon formal specification and constraints.”

The proposed combination of terminological principles (onomasiologic approach) with the lexicography (semasiologic approach) is not novel and discussed from a term formation point of view by Fernández-Silva, Freixa & Cabré (2011, p. 54) stating that “special knowledge is located in the experts' minds and the commonest way of approaching concepts and conceptual structures is by means of their linguistic representations, i.e., by analyzing specialized texts.

Constructing dual entries

Following Rogers (2004), who attempts to construct concept systems using textual cues, we are able to pair a term with its underlying concepts. For instance, “energy tax” (*energiskat*) is presented in a bilingual article format with a formal definition and assigned to the underlying system of concepts with the designated concepts in the centre of the diagram (see figure 4.3). We may assign the concept diagram, which graphically displays the terminological ontology of the particular concepts and closely related concepts reusing the terms as labels on entities and formal feature specifications (Freixa, 2006), to the article entry displaying written categories of data about the term. The information format connecting the concepts and terms is called dual-entry mode.

It should be noted that the proposed dual-entry modes must also comply with the balanced requirements outlined in the so-called lexicographic triangle of Verlinde, Leroyer & Binon (2009), where the “triangulation of data, user and access” or the “user-adapted access to data “ensures the performance of the tool. But a fourth interaction dimension is needed for the actual fulfillment of the lexicographic function, which emphasizes logical access points and search routes, leaving no room for random access, dead ends, lucky guesses or heureux hasards, no room in other words for serendipity" (Verlinde, Leroyer & Binon, 2009, p. 5) The underlying ontologies supports graph-based access routes, i.e. target users commence their search by navigating the terminological ontology.

2.4 Conclusion

In the semantic view on domain-specific terminology, it can be concluded that the study of terms rests on three theoretical criteria regarding the discourse, the lexicalization and the underlying concepts. However, the framework is vague as it does not ensure the identification of all terms, which differ in distance to natural language and degree of domain specificity (see figure 2.1) potentially concealing the terms. I conclude that the criteria can easily be questioned with border case examples, but in the chosen taxation domain, I am able to proceed with a core set of domain-specific terms on which I conduct the terminology work and construct the terminological ontology.

I conclude that terminology work is iterative allowing for a combination of the semasiologic with the onomasiologic approach. The process begins with the extraction of potential terms from the discourse documents, clarifying the underlying concepts to build terminological ontologies, which are formally describing a domain of discourse (potentially supporting computer applications), by interpreting the semantic contents and deriving definitions from the consistent conceptual structures. The process may be repeated in circles until outstanding terms or concepts are clarified.

Term banks are constructed to provide users with the data resulting from the terminology working process. I conclude that both the conceptual and lexical sides of terms in combination with the authentic specialized discourse i.e. terms should be accessible to target users of term banks. Often term banks follow the term-oriented specialized dictionaries in displaying data (e.g. term, source, definitions, equivalents etc.). I conclude that terms may be seen as devices to access concepts of an ontology, which allows for the combination of functional lexicography emphasizing the needs of particular user situations with formal ontology renders possible dual-entry modes of the term-bank user interface. Here term banks visualize the concept-oriented terminological ontologies (based on graph theory) and assign them to each of the term-oriented textual entry (based on lexicographic function theory) in dual-mode entries.

Regarding research themes, I conclude that dual-entry modes displaying the term and concept pairs provide much more accessible information, potentially enhancing knowledge dissemination to target users.

Chapter 3: Terminology and knowledge

Chapter 3 is central to analyzing the second research theme uncovering multiple domain-specific cultures term variation and turning underlying target-user expertise complex. In addition, the relevant aspects of domain-specific knowledge and knowledge acquisition shaping target user behaviour.

As will become clear, specialized communication aims at conveying knowledge across cultures. In the case of domain-specific knowledge, the motivation for term variation rests on the need to cross barriers of domain cultures shared by what we may call communities of practice, communities of knowledge and communities of language. In other words, the term variation is motivated by the need to convey knowledge across cultures of practice (the so-called intra-domain culture in section 3.2), knowledge cultures (the so-called extra-domain culture in section 3.3) or language cultures (the so-called supra-domain culture in section 3.4).

The pragmatic view on terminology uncovers the potential barriers that target users face when disseminating or acquiring domain-specific knowledge, which provides us with key aspects of the development of the term bank: First, regarding the content, the framework is based on term variation and will therefore help uncover the terminological landscape (i.e. terms, synonyms, homonyms and equivalents) of a particular domain, which should be represented in the term bank. Second, regarding the target users, the framework allows us to model typical users into a matrix, which uncovers the relevant dimensions of user adaption as means to ensure the usability of the term bank. Third, regarding users' performance in an experimental setting, when searching the dual-entry mode (developed in chapter 2), the framework uncovers the complexity of the domain-specific expertise.

I review the knowledge about cross-cultural communication to introduce the notion of domain culture. Culture has received extensive attention in the communication literature, but I focus on those aspects which are particularly relevant to term variation (section 3.1). In particular, I examine the motivations for term variation along three dimensions: The intra-domain cultures (section 3.2) followed by the extra-domain cultures (section 3.3) and finalized by the supra-domain cultures (section 3.4). I develop a graphic illustration containing the dimensions of domain culture in theory and practice (section 3.5). Finally, the domain-specific knowledge is separated into underlying types of knowledge, which are important for explaining term-bank user behaviour and performance (section 3.6).

3.1 Introduction

A vast body of literature has shown that people acquire knowledge at different rates according to differences in aptitude. (General) problem solving aptitude has the largest effect on (procedural) knowledge acquisition (Bonner & Walker, 1994). Mason & Singh (2011) suggest that users should be located on a multidimensional continuum, where categorization is one aspect of expertise, instead of the expert and novice labels. Learners with low prior knowledge do not use multiple representations (Seufert, 2003). Learning topics in science is a matter of altering prior conceptions (conceptual change), and adding new knowledge (Rusanen & Pöyhonen, 2013). Conati & Merten (2007) use eye tracking to show meta-cognitive behaviour of adaptive applications.

Friege & Lind (2006) discuss different kinds of knowledge underlying problem solving. No commonly accepted cognitive-psychology definition of “expert” and “novice”, but the former is a domain-specific phenomenon presupposing extensive practice. Friege & Lind (2006) underlines that expertise is relative, i.e. in one group a person can be classified expert, whereas in another group he will be regarded novice. Essential differences lie with the type of knowledge (declarative, procedural and problem-solving) and the storage (i.e. long-term versus working memory)

Graves (1996) characterizes expert performance by means of the required knowledge of a specific domain as well as the appropriate strategies (domain-specific expertise). It follows from this framework that the study of novices reflects general strategies to be used, when the appropriate knowledge is missing (generic expertise). If domain-specific experts are facing demanding (unfamiliar) problems in their domain, they (also) have to rely on more general strategies.

Britton, Stimson, Stennett, and Gülgöz (1998) (cited by Cuevas, Fiore & Oser, 2002) discussed four variables affecting knowledge acquisition (i.e. metacognition, inference-making ability, working memory, and domain knowledge. Meta-comprehension refers to the "conscious processes of knowing about comprehending and knowing how to comprehend" (Osman & Hannafin, 1992, p. 85, cited by Cuevas, Fiore & Oser, 2002)

Domain-specific cultures

In the pragmatic view on domain-specific terminology, I emphasize the context surrounding terminology users, which I approach from a cultural point of view, since culture shapes our view of the world and our thinking. However, the term culture is widely used especially in the fields of psychology, management and communication, and I do not intend to offer a universal definition. Still, there are good reasons for sticking with the broad and multifaceted term culture, despite the risk of misleading some of my readers. Firstly, I interpret the renowned cultural dimensions theory proposed by Hofstede (1984), which is a framework for indexing key factors of cross-cultural communication, as I propose three dimensions of domain culture shaping domain-specific terminology by motivating term variation. Secondly, I apply and extend the conventional (one-dimensional) culture-bound attribute that scholars of translation studies assign to terms, which are subject to nonequivalence between SL and TLs. In domain-specific terminology, the culture-boundness of terms is multi-dimensional, and therefore I prefer the "culture-specific" side of terminology over the narrow "culture-bound."

One-dimensional target-user expertise

Functional (i.e. user-oriented) lexicographic theory aims at meeting the complex of needs of users in particular situations. Concerning user expertise, users are traditionally organized on a one-dimensional scale from experts to semi-experts and laymen (Caruso, 2011, citing Bergenholtz & Kaufmann, 1997), where experts are characterized by relying on other sources than term banks or dictionaries to acquire new knowledge. This is necessarily not the case with semi-experts, who are characterized by their near relation to a subject field, like for instance journalists who disseminate issues from other fields or advisors in public administration becoming familiar with the sectors of their authority. The laymen are characterized by having little prior knowledge about a subject field, but as opposed to the expert and semi-experts, he is very likely to search the internet (including term banks) for knowledge. However, we need to both engage the domain experts in the building of the system, and we need fruitful ways of attracting them to the tool as target users. Studer, Benjamins & Fensel (1998) outlines development of first generation knowledge banks emphasizing the transfer of knowledge from expert to base, while the second generation knowledge banks, which are the result of modelling activities simulating the problem-solving capabilities of a domain-expert (expert systems). But as we shall see, term banks offer little in the case of domain-specific problem solving (see table 3.3).

Three-dimensional target-user expertise

Seen from the view of terminology theory, three cultural dimensions appear: It is assumed that terms are not invented but exist in a specialized discourse, where they constitute lexical units and designate an underlying concept. An expert in domain-specific terminology and knowledge is not merely consisting of specialists of particular sciences (e.g. law and finance), I also encounter the LSP communications specialists who master the conversation (in SL) with the entire communities of practice (but without challenging the conceptual knowledge like the field specialists), as well as the LSP translation specialist who solves the equivalence problems in TLs with appropriate translation strategies. Moreover, it is likely that a subject field expert of taxation cannot account for the taxation systems of other countries or regions, and he will exhibit low LSP translation expertise. All three dimensions (practice, knowledge and language) are necessary to fully account for the complex views on terminology, which shapes the needs of potential target users of the term bank. All three dimensions are necessary for developing sound measures of expertise. In particular, I develop a graphic illustration of the framework (see figure 3.1 in section 3.5), where I elaborate the discussion of target users and their context (pragmatics) in the chosen taxation domain, which should be supported by the term bank.

It should be noted that the framework rests on two critical assumptions: First, I assume that a domain-culture exist, i.e. a set of terms including the underlying concepts are shared by a group of people. Gruber (1995, p. 909) describes it as ontological commitment, i.e. “[p]ragmatically, a common ontology defines the vocabulary with which queries and assertions are exchanged among agents. Ontological commitments are agreements to use the shared vocabulary in a coherent and consistent manner.” Second, I assume that a motivation for term variation develops from the need to disseminate knowledge. Following their work on cognitive dynamics, Fernández-Silva, Freixa & Cabré (2011) introduces term variation by asserting that “concepts and conceptual structures adapt to the speakers' cultural, social and situational environment. One of the clearest manifestations of these cognitive dynamics is the lexicalization of concepts through various expressions, or terminological variation.” In addition, term variation is motivated by the need to convey a particular vision to the target audience. Fernández-Silva, Freixa & Cabré (2011, p. 52) states that “[a] concept can be expressed by a single term or by several terms that convey the same meaning” but “each term displays a particular vision of the concept (Freixa 2006).” I show an example of the term variation in the chosen taxation domain in figure 3.2.

3.2 Intra-domain cultures

In this section, I discuss the proposed intra-domain cultures shared by communities of practice, where term variation (among synonyms) is motivated by the need to cross levels of specialization.

The label “intra” means inside (Oxford Dictionaries, 2015). I have borrowed the label intra-domain from Al-Sayed & Ahmad (2006), who uses it to modify the registers (i.e. intra-domain registers) shaping the terminology inside the communities of practice in cancer care. According to Al-Sayed & Ahmad (2006) communities of practice exist within a domain with their own distinct exclusive terms, which constitute common vocabulary the so-called communal lexicon. In addition, that in distinguishing domain from discipline, Alexander (1992, p. 36) accounts for so-called internal criteria as the extensiveness of individual's knowledge.

In her typology over causes for term variation, Freixa (2006) outlines a five-dimensional approach to term variation, which I collapse into three relevant dimensions in the case of domain-specific terminology. The functional motivation is merged with the discursive (i.e. pertaining to stylistic needs), the linguistic motivation is merged with the dialectic (i.e. pertaining to author origins) motivation, while the cognitive one is maintained. So the intra-domain cultures encompass cultures of practice in a domain, where the term variation is motivated by the functional need to convey knowledge (in the SL) (see section 3.2.1), we return to the linguistic (see section 3.4) and cognitive (see section 3.3) below.

3.2.1 Communities of practice

The intra-domain culture shared by communities of practice motivates the term variation to ensure effective expert-to-non-expert knowledge dissemination. Language users are separated into high, medium and low levels of specialization: Al-Sayed & Ahmad (2006) conduct a case study of cancer care, with three levels including researcher, practitioner and patient and proposes to use terminology sharing as a metric for knowledge sharing. Often we encounter the levels experts to semi-experts and laymen (see e.g. Caruso, 2011 and Da-Cunha et al., 2011). Finally, Cabré (2003, p. 188) underlines that the specialized discourse contains a number of communicative scenarios, for instance, communication among specialists, between specialists and semi-specialists or technicians, between specialists and learners."

As mentioned, term variation is motivated by the functional need to disseminate knowledge across the communities of practice, but adhering to the same conceptual view (fundamental concepts).

For the sake of argument, I prefer to follow the conventional three-level approach in the taxation domain, however, this discrete variable of expertise may also be conceived in a continuous scale (see figure 3.2): researchers (with high expertise) develop knowledge in high volumes and in abstract forms (e.g. indirect consumption tax), which have the potential of being disseminated to professionals or practitioners who have to understand, critique and apply the knowledge developed (medium expertise), while clients and patients (low expertise) are targeted by the two former by using less abstract terminology (e.g. value-added tax).

3.2.2 Examples of specialization

As stated by Andreoni, Erard & Feinstein (1998, p. 818): “[t]he economics of tax compliance can be approached from many perspectives: it can be viewed as a problem of public finance, law enforcement, organizational design, labor supply, or ethics, or a combination of all of these.” Therefore, some areas will inevitably be engaging more intra-domain players than others.

Researchers of taxation (or any field for that matter) may primarily be concerned with their knowledge production, i.e. the knowledge dissemination to the community about the conceptualizations governing their field. However, inside the intra-domain culture, I emphasize the communication of a given concept to lower levels of abstraction (non-experts). We may choose the example of researchers (high level specialization) who have demonstrated the welfare effects of reducing the distortions of commodity taxation (see table 3.1). These results should be conveyed to teachers and to students of the field. Moreover, to realize the welfare gains of the mentioned example, legislators (high level) may be induced to act, and practitioners (medium level) in the field must be informed about the potential, otherwise they are not able to suggest and draft the legislation necessary. Finally, welfare gains presume taxpayer compliance, which is provided for through legal instructions. The different levels of specialization employ different levels of abstraction: For instance, indirect consumption tax, value-added tax and increased sales tax (see figure 3.2). The tax professionals (practitioners) need to explain the taxation terms used.

Depth of expertise

The chosen direction for the level of specialization follows the employed horizontal and vertical axes by Da Cunha, Cabré, SanJuan, Sierra, Torres-Moreno & Vivaldi (2011, pp. 266-267) stating questioning the notion of a specialized text: “[t]here are two types of variability in specialized texts: horizontal determined by the subject and vertical determined by the specialization level. With regard to the second one [...], three specialization levels can be considered: high (specialized writer and specialized receiver), medium (specialized writer and semi-specialized receiver, that is, for example, students) and low (specialized writer and non-specialized receiver, that is, general public)." Moreover, the (horizontal) breadth and (vertical) depth are also employed by Verlinde, Leroyer & Binon (2009).

We may confirm that intra-domain culture is a question of depth (vertical direction in the terminological ontology): The Danish tax system is shown and the terms of the different practices cluster around fundamental concepts (see section 3.3). Following the argument stated above we may navigate the terminological ontology vertically: In the bottom of the terminological ontology, the individual taxes (e.g. wine duty) are depicted and here the source material is the tax laws which all taxpayers are liable to and where the taxpayer compliance is accounted for. In the middle, consolidated or general concepts are constructed to facilitate the authorities' management of tax compliance (e.g. excise duties). Finally, looking in the top of the terminological ontology, we see a whole different set of terms extracted (e.g. indirect taxation) from a completely different kind of expert discourse, namely economic theory, where the level of abstraction is much higher, devoted to economics and public financial management.

Target-user matrix

I introduce the notion of a target-user matrix, see table 3.1, which resembles a user model of the term bank accounting for two of the three dimensions (i.e. depth and breadth, but not height). As mentioned in the introduction, the proposed framework uncovers the potential cultures and barriers underlying the acquisition of knowledge, which reflect the user needs that the term bank must meet and I use the framework to outline examples of potential user groups to be targeted by the term bank.

Table 3.1: Target-user matrix. The matrix shows examples of (stereotype) target users ordered by their level of specialization (high, medium and low) and adherence to disciplines of taxation (tax policy, tax compliance, tax avoidance, taxation theory). For the latter, the degree of domain-specificity is indicated going from the less specific (general) to the highly specific (specific). The matrix disregards the target group of translators. The terminological ontology of appendix II follows the target users assigned to the tax-collection column.

	Public finance	Revenue collection	Tax avoidance	Taxation theory
High	*Politicians*	*Legislators*	*Tax planners*	*Researchers*
Medium	*Journalist*	*Practitioner*	*Accountants*	*Teachers*
Low	*Newspaper readers*	*Taxpayers*	*Clients*	*Students*

The table shows three levels of specialization and four different knowledge communities. The latter is discussed in the following section.

3.3 Extra-domain cultures

In this section, I discuss the proposed extra-domain cultures shared by scientific communities, where term variation (among homonyms) is motivated by the need to cross different conceptions of the field.

The label “extra” means outside (Oxford Dictionaries, 2015). I have developed the label extra-domain from the term extra-linguistic, which L'Homme (2003) uses to emphasize the extra side of terminology beyond the pure linguistic, i.e. the lexical units that terms contain have an underlying relationship with a given subject field. Roche (2006) asserts the missing direct relation between ontology (i.e. conceptualization in the form of extra-linguistic representation of conceptual knowledge) and text (linguistic matter). We also encounter the term extra as a modifier of the lexicographical situation (Bergenholtz and Tarp, 2010).

In her typology of causes for term variation, Freixa (2006) argues that different conceptualizations require the cognitive motivation for term variation. So the extra-domain cultures encompass scientific cultures of knowledge (i.e. communities building knowledge and forming concepts) in a domain, where the term variation is motivated by the conceptual need to convey knowledge (in the SL). In addition, Fernández-Silva, Freixa & Cabré (2011, pp. 53-54) conclude on the basis of their literature review "experts make conceptually motivated term choices", or put differently, "experts from different subject fields make different term choices when referring to the same concepts." This is the case for the different conceptualizations underlying e.g. value-added tax, which may implicitly be the act of imposing the tax, the rate imposed, the legal authority underlying the imposition etc. (see figure 3.2).

3.3.1 Communities of knowledge

Fernández-Silva, Freixa & Cabré (2011, p. 51) states that "[t]erms, the linguistic representations of concepts, are created by a particular language community which needs to communicate the knowledge it has produced. In addition, "[t]he structure of knowledge can change for two reasons: a new axis may appear (redefine concepts wrt new dimension) [...] or a new way of seeing things may arise (position change)," (Sager, 1990, cited by Cabré (1999, p. 43). This leads us to believe that extra-domain cultures imply scientific barriers impeding the communication between experts (expert-to-expert communication).

At the surface it may seem that the framework violate the universal view on concepts, as Cabré (2003, p. 167) states that "a concept is universal[1], independent of cultural differences and that consequently the only variation possible is that given by the diversity of languages." However, I am not challenging the notion of concepts per se, instead I emphasize the different views on conceptualizations (i.e. concepts, characteristics and relations to super-, side- or sub-ordinate concepts). Cabré (2003, p. 179) states quite clearly that "in many subjects there is no unified conception of the field." In particular, I employ the cognitive dynamics of building new knowledge by introducing the extra-domain culture. Following Cabré (2003, p. 166) saying that terminology is basically "the semiotic conception of designations," I apply the study of signs processes triggering the cognitive development of concepts described by Thellefsen & Thellefsen (2004), which result in the linguistic realization. In other words, "[t]erms are also the final step in concept formation," (Fernández-Silva, Freixa & Cabré, 2011, p. 51)

[1] To some researchers, concepts are language dependent

Fundamental signs

Concepts reflect the way players of the knowledge domain organize their knowledge by reducing the complexity of a so-called knowledge potential (Thellefsen & Thellefsen, 2004). The (stable) interpretation (practice) of a particular concept is highly controlled by the specialized discourse establishing a pre-understanding, which defines the meaning potential of the concept in accordance with the given specialized context (Thellefsen & Thellefsen, 2004). Consequently, the actual knowledge release from a particular concept will depend on the knowledge volume of the interpreting individual, manifesting parts of the concepts' knowledge potential in the players' concrete use of the concept (Thellefsen & Thellefsen, 2004).

The same concept may exist in different knowledge domains, thereby creating a meaning potential reaching beyond the pure terminological level (Thellefsen & Thellefsen, 2004), i.e. the taxation concepts exist in several domains e.g. economics, law and political science. A domain player from inside the knowledge domain will have larger knowledge about the concepts originating from that particular domain, than players from outside the domain (Thellefsen & Thellefsen, 2004), i.e. economists will know about their concepts, which the legal players will know little of and vice versa.

Over time a common, general idea is maintained, which creates the background for the development of the knowledge domains so-called fundamental sign, from which the majority of the domain's knowledge is organized, constituting the terminology of the domain (Thellefsen & Thellefsen, 2004). The fundamental sign expresses the knowledge that the majority of players in the knowledge domain agrees on is the focal point of the field (Thellefsen & Thellefsen, 2004), i.e. the fundamental sign is the result of a democratic process and leaves perhaps a minority not affiliated with the idea.

Some concepts contain a larger knowledge potential than others, because they have a more significant meaning to the knowledge domain players than other concepts (Thellefsen & Thellefsen, 2004). The fundamental sign is the sign with a large number of related concepts and is the reference point in the conceptual structure of related concepts (Thellefsen & Thellefsen, 2004) e.g. “punktafgift” (see figure II.1 in appendix II). At the same time, it should be noted that a peripheral concept can be a fundamental sign of a different knowledge domain (Thellefsen & Thellefsen, 2004), i.e. concepts far from the financial fundamental signs may constitute fundamental signs of law or political science. It is the development of the fundamental sign that creates the basis for the development of the knowledge domain's specialized language (Thellefsen & Thellefsen, 2004). Each new concept is adapted into the exisiting knowledge structures. Knowledge is developed around their fundamental sign, while for the players outside the domain (external), this means that knowledge is based on a fundamental sign different (or far) from their fundamental sign(s).

3.3.2 Examples of fundamental concepts

In this section, I give examples from the taxation domain illustrating the fundamental concepts. In my case study, I assume that experts share the same SL (Danish), but approach the taxation domain differently from each of their disciplines. The purpose of this type of communication could be to improve the efficiency of the national tax administration e.g. preparing a minister for meetings with his fellow ministers, internal communication between legal staff members and economists communicating to reach an estimate of the financial effects of a tax proposal.

Breadth of expertise

In the terminological ontology (see figure II.1 in appendix II), we realize that extra-domain culture is a question of breadth (horizontal direction) as opposed to the vertical direction of intra-domain cultures. However, the terminological ontology lies within one single conception of revenue collection (see tabel 3.1) based on the constituent taxes and duties. Therefore, I am not directly questioning any conceptual issues at the scientific level. But it is not hard to imagine a complete mapping of terminology of the taxation domain, where for instance green taxation (see figure D.5 in the appendix) based on modern economic theory does not fit the established revenue structures.

In the target-user matrix (see table 3.1) along the extra-domain dimension (horizontal): I have stated four examples covering the less specific tax policy over revenue collection, tax avoidance, and the highly specific taxation theory. Differences in interpretation practice trigger sign processes, which result in the construction of knowledge. In the taxation domain, researchers of taxation theory may (again) demonstrate welfare gains of reducing taxation, but domain players of compliance or avoidance have different interpretation practices of the fundamental concepts.

3.4 Supra-domain cultures

In this section, I discuss the proposed supra-domain cultures shared by linguistic communities, where term variation (among equivalents) is motivated by the need to overcome nonequivalence between languages. Much terminology work is devoted to translation (Nkwenti-Azeh, 1994).

The label “supra” means beyond (Oxford Dictionaries, 2015). I have developed the label supra-domain from the term supra-national, which is widely used by the European Union, which in some cases is given supra-national authority.

In her typology of causes for term variation, Freixa (2006) argues that different languages may require the linguistic motivation for term variation. So the supra-domain cultures encompass language cultures in a domain, where the term variation is motivated by the linguistic need to convey knowledge in the TL.

3.4.1 Communities of language

I emphasize that supra-domain cultures establish linguistic barriers impeding the translation at the national, regional or global levels. Historically, terminology and term banks have been closely linked to the field of translation (Nkwenti-Azeh, 1994).

It should be noted that in international communication, we often talk about a language as belonging to a nation. But in the case of domain-specific terminology, it is important to determine the territorial boundaries. In the choice of domain the legal territory (jurisdiction) is the dominating one, which means that the (tax) jurisdictions around the world are establishing the terminology. In practice, the jurisdictions are both horizontally shaped by national borders (e.g. Denmark) i.e. national languages form the basis for the basic terminology. However, the tax jurisdiction may also be organized vertically in that governments at different levels (both national and regional) may have the authority to impose taxes on the same territory (e.g. Denmark is a member of the European customs union) and even the same tax base, giving rise to double taxation and tax competition. At this level, a parallel terminology in both Danish and EU English (different from the British English tax terminology covering the British territory) will be the result. Finally, global organizations are highly influential forums without direct taxing authority, where governments share experience and work together to promote efficient tax policies around the world. For instance the OECD may shape the international terminology of the taxation domain by introducing conceptualizations necessary for encompassing the international community in the same framework of analysis from which recommendations are inferred (e.g. value-added tax, goods and services tax and general sales tax, see figure 3.2). It should be noted that in the taxation domain, supra-domain culture is not only a matter of country descriptions, since a country is simultaneously part of several cultures apart from the national level, a country is part of a region (e.g. the EU), and part of the global economy (e.g. the OECD).

Equivalence across languages

Bilingual dictionaries accounts for equivalence. Sandrini (1999) ascertains that "[t]raditionally, legal dictionaries on the market are the product of a lexicographical approach to legal terminology listing different meanings of one word and proposing possible equivalents in the other language." Moreover, Adamska-Sałaciak (2010) states that "[t]he relationship between words or phrases, from two or more languages, which share the same meaning." i.e. "[i]n bilingual or multilingual terminological dictionaries, equivalence implies interlingual correspondence of designations for identical concepts." In addition, the problem of inequivalence is discussed by Adamska-Sałaciak (2010) stating that "[b]ecause of linguistic and cultural anisomorphism, translation equivalents are typically partial, approximative, non-literal and asymmetrical (rather than full, direct, word-for-word and bidirectional)."

Sandrini (1999) abandons the notion of equivalence in favour of a comparative approach i.e. “we have to abandon the concept of equivalence in favour of a more flexible comparative approach. The difference lies in the pre-supposition that legal concepts as part of a national system of laws are fundamentally different across legal systems and that only a comparative approach is possible”. In the choice of domain (taxation), we need to consider the regulatory intention, i.e. “The basis for comparison is the function of each concept within the legal setting expressed by a functional definition which describes the role of the concept with regard to the overall regulatory intention of the whole legal setting.”

“Such a comparative approach is not intended to lead to one-to-one equivalents in a dictionary or glossary. Its main goal is to convey information on the concepts within each legal system and to offer some kind of bridges between the two legal systems in order to lead the user from one concept in one legal system to comparable concepts in the other legal system.

Translation strategies

In table 3.2, I outline possible translation strategies across different levels of equivalence, where effective solutions range from lexicalization or term-formation by adopting a foreign word or producing a direct translation to descriptions by translating the definition or producing a vague, but idiomatic explanation. The consequences of the translation strategy for the SL domain are outlined in table 3.2.

Table 3.2: Translation strategies across levels of equivalence. In the case of true (natural) equivalence, direct translations are possible and the domain is preserved. In the case of near or nonequivalence (directional equivalence), no natural equivalents exist and second best approaches must be employed to form terms or phrases. False equivalence should be avoided.

	False friends (single terms) (1)	**False friends (generalized terms) (2)**	**True equiva-lence (3)**	**Near equivalence (4)**		**Nonequivalence (5)**	
Strategy	Avoid direct translation	Avoid general designation	True mapping: Direct translation	Adopt the foreign term directly	Literal translation	Use nearest equivalent	Definition or explana-tion
Example	*staten ≠ state government = central government*[2]	*grøn afgift ≠ green tax; boligskat ≠ home tax; pensions-skat ≠pension tax*	*CO2-afgift = duty on CO2; direkte skat = direct tax*	*transfer pricing; com-pliance*	*tax gap = skattegab; fiscal tax = fiskal skat*	*duty on coffee; motor vehicles registra-tion duty*	*virksom-hedsskat = tax imposed under the Coporate Tax Act as a prepaid tax on accumulate d earnings*
Effect	Cross-linguistic contamination (⟐)		Domain preser-vation	Domain Loss	Domain preser-vation, but idiomatic erosion	Term preser-vation when reversing translation direction	Term loss when reversing translating direction

3.4.2 Examples of translation

In this section, I give examples from the taxation domain illustrating the translation strategies to overcome the degrees of nonequivalence, outlined in table 3.2.

[2] According to EU; Germany has both central and state taxes.

In the taxation domain, terms are closely attached to the foundation and structure of the Danish legal system in which the taxation laws apply. At the general level many taxation terms have true equivalents in many foreign languages e.g. “carbon dioxide tax” (*CO2-afgift*) equivalent to duty on CO2 (carbondioxide) (see column 3 of table 3.2), which is one of the focal instruments to be discussed when the world gathers for climate change discussions (e.g. limiting greenhouse gas emissions agreed to in the Kyoto Protocol). In this case, there is full equivalence, since the CO2 is a generic substance (physics) and the purpose is to lower the emissions in order to limit the harmful effects of consumption. However, in the taxation domain many other especially societal factors or practices come into play, which erodes the equivalence, e.g. the Danish 'state” (*staten*) is not equivalent to the 'state government” in Germany (rather we need to turn the focus to the central government) (see column 1 of table 3.2). Below I discuss solutions to inequivalence.

False equivalence (false friends)

When translators encounter false equivalences the only reasonable strategy to pursue is to avoid the direct translation, which exist in the target language but is of a different meaning. Instead concept clarification should reveal true equivalents or explanations and maybe even a note elaborating on the features of the false equivalence. False equivalence is potentially very damaging to the specialized discourse, since it might very well be impossible for target readers to discover the mistaken use of the term, since the lexical unit will exist.

If we consider true and false equivalence of general terms (see column 2 of table 3.2): In the taxation domain, we are often discussing taxation structures (as part of fiscal consolidation) i.e. the distribution of tax revenue across the various types of taxation and therefore it is useful to introduce new generalized and popular terms and therefore useful in the public debate. Generalized terms may not necessarily constitute a problem e.g. “direct tax” (*direkte skat*) is equivalent to direct taxes. However, in many cases, and especially if certain country-specific groups of taxes (or a system of high granularity as the Danish tax system) are aggregated, problems of false equivalence arise. One example is the Danish “green duties' (*grønne afgifter*) which is translated into green taxation by the EU. In Denmark, the Ministry of Taxation includes the energy taxes, environmental taxes and motor vehicles taxes including in particular the registration tax in the green taxation category, despite the fact that the latter has very little to do with taxes designed to prevent pollution or protect the environment. In fact, also “green” cars are taxed. So it is highly questionable whether the collective terms constitute equivalents at all, and whether direct translations distort the figures and debate with approximate measures, unless a clear definition underlining differences is included.

Near equivalence

Let us consider the case of near equivalence (see column 4 of table 3.2). It is very common to come across near equivalence which in domain specific contexts or with special or distinct language may not constitute any problems since no confusion is made. Still the translator should be very delicate about this task, otherwise the translations will have a negative impact in terms of domain loss or idiomatic erosion. One way of overcoming near equivalence is to lexicalize the term i.e. produce a literal translation in the target language (glossing) e.g. “tax gap” was translated into *skattegab* and “fiscal tax” has been translated into *fiskal skat*. The consequence of this term formation is domain preservation, since the lexicalization contributes with proposing new members of the specialized lexicon.

However, there is a risk of idiomatic erosion since the terms that are being lexicalized may not make much sense to laymen. The result is a distorted discourse, i.e. terms are formed and used in texts which are not adopted by the expert community, though this is not the case with the examples given. Another way of solving near equivalence is to adopt the foreign term directly into the target language e.g. “*transfer pricing regler*” (rules) and “*compliance undersøgelse*” (investigation) are used in the Danish discourse, resulting in a confusing mixture of different languages. The consequence is domain loss since foreign terms crowd out potential Danish equivalents. In the examples given, Danish equivalents do exist (*fastsættelse af interne afregningspriser* and *regelefterlevelse*, respectively), but they were never adopted in the discourse where the English (OECD) term was preferred over the Danish lexicalization.

Nonequivalence

Let us consider nonequivalence (see column 5 of table 3.2): As we saw above, taxation terms are linguistic variations over titles from the body of Danish legal discourse (apart from being pure revenue collection tools), which by nature is Danish, and that constitutes the reason of nonequivalence, in the same way that titles produce nonequivalence. Strictly speaking, one could argue that the legal language in Denmark is Danish, which implies that all the conditions of the law should be reflected in the target country, otherwise nonequivalence will be the result. In this case, the terms are not existing in the target language, where the corresponding terminological ontology will be simpler covering fewer taxes. This challenges the translation task differently, since we no longer need to avoid the risk of choosing false equivalents, rather, we need to form a term that makes the target reader grasp the content of the term (and concept) without being familiar with the term. If the terms are straightforward, the translation strategy is to begin with the nearest common super-ordinate equivalent, possibly duty, e.g. (*kaffeafgift*) is translated into "duty on coffee" or (*registreringsafgift*) is translated into "motor vehicle registration duty", which preserves the terminology in both source and target texts when the translation direction is reversed.

However, some cases arise, especially if there is a large distance between the source and target languages, where conceptual structures do not match because of high complexity and then the definition is the only available point of departure for the translation strategy, e.g. "top-bracket tax" (*topskat*) will not make sense to anyone not familiar with the required progressive, state income taxation system with at least two income tax thresholds. Maybe this is why the OECD defrays from using specific terms, however, the problem is that when the translation direction is reversed the term is lost. Translating by producing a vague, but idiomatic explanation e.g. replacing "*virksomhedsskat*" with "*skat som iht. Virksomhedsskatteloven opkræves som en acontoskat på det opsparede overskud*" (embedding terms in TL) is not advisable, since the consequences of reversing the translation direction is that a term disappears.

Standardisation

Wüster's motivation came from an increasing international trade in industrial products which could be eased by an international standardization of terms used in engineering ensuring unambiguous specialized communication in international contexts. In the modern knowledge-based economy, globalization increases the mobility of tax bases, and the various taxation rules constitute barriers to the free movement of goods and services, and stronger cooperation at governmental level is called for to avoid horizontal tax competition. But maybe there is a renewed need for standardization to improve trade and communication of the knowledge economy, explaining taxpayers (individuals and companies) key features about taxation which are otherwise inaccessible or too time consuming to understand.

Compliance is closely linked to taxpayers' trust in the authorities, which is likely to increase if authorities adhere to international standards of tax administration (Boll, 2011) confirming the point made by Kleven, Knudsen, Kreiner, Pedersen & Saez (2011) that tax compliance depends on psychological or cultural (non-economic) aspects. Standardization is already a reality at the supranational regional level in the EU, e.g. national accounts could be the point of departure, which could be the basis for a global standard at OECD level.

In the taxation domain, a well-established international discourse exists due to the supra-national authorities assigned to the EU (legislative authority) and the OECD (indirect authority of recommendation), the forums of the EU discussing harmonization, harmful national tax practices resulting in tax competition and tax evasion. But on a global scale, the OECD, more countries are included, which puts even stronger demands on the lingua franca to encompass all details. The multilingual EU discourse offers a bilingual (Danish-English) publication, which is not the case for the OECD, which only communicates in limited official lingua francas (English and French).

In the terminological ontology (see appendix II), we realize that supra-domain culture is a question of exhibiting a third (non-displayed) dimension opposed to the breadth (horizontal direction) of knowledge and the depth (vertical direction) of practice. (In practice, the TL information is accessed by clicking the individual nodes of the system navigating to the underlying entry displaying written terminological information). In the figures 3.1 and 3.2, the supra-domain culture is realized as height.

3.5 Graphic illustrations of the domain-culture framework

In this section, I propose a graphic illustration displaying the dimensions (depth, breadth and height) of the cultures of domain-specific terminology (see figure 3.1), which allows us to examine relevant examples (see figure 3.2).

The multi-dimensionality is illustrated by means of the co-ordinate system comprising axes in three directions (x, y and z) and inhabitating the space with term variants around a fundamental concept (origo). This approach is inspired by Cabré (1999, p. 43) describing: "Sager (1990) [who] presents a model of knowledge in the form of a multi-dimensional space made up of a series of intersecting axes, each of which represents a class of conceptual characteristics or dimensions in an intersecting relationship."

The intra-domain culture shared by communities of specialization is realized along the y-axis and is a matter of depth, while the extra-domain culture shared by communities of knowledge are realized partly along the x-axis and are a matter of breadth. The supra-domain cultures shared by communities of language is realized along th z-axis and is a matter of height. The illustration allows for the location of SL terms with z=0, while TL terms will locate themselves for z>0.

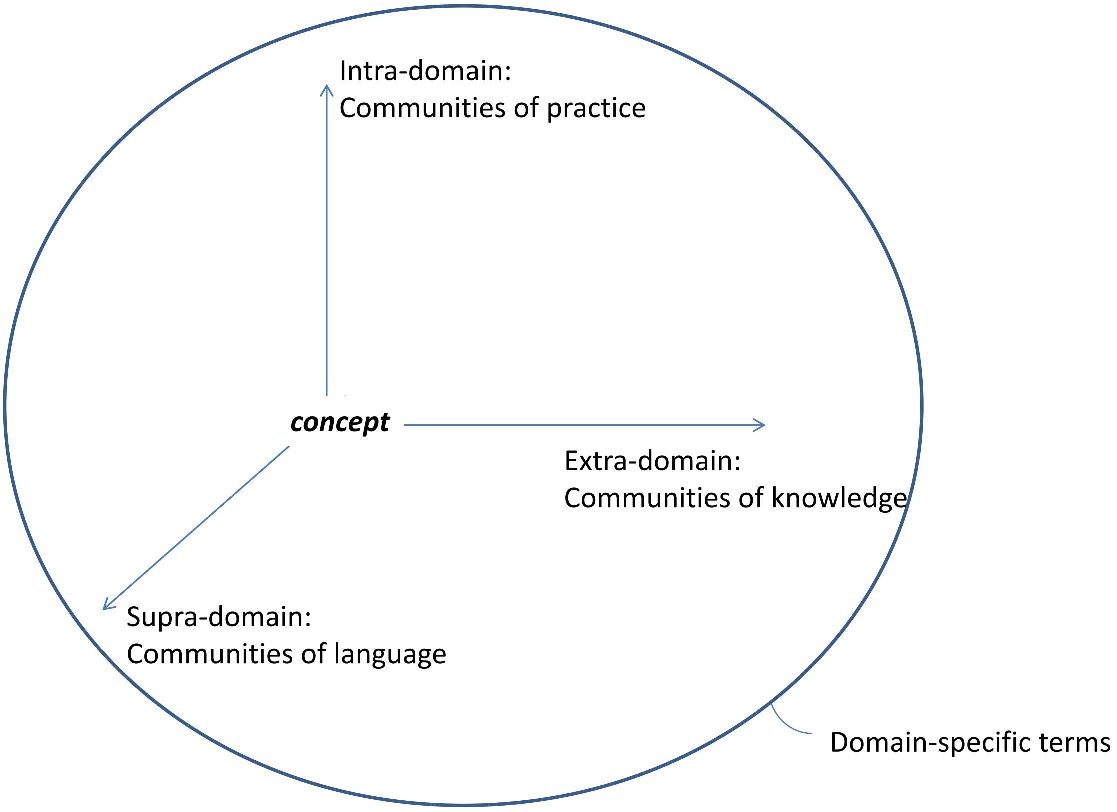


Figure 3.1: The motivation for term variation in communities of knowledge, practice and language. In the SL (z=0): Intra-domain cultures shared by communities of practice: Given a concept (meaning), synonym terms are chosen to convey knowledge across different levels of abstraction. Extra-domain cultures shared by communities of knowledge: Given a term, different concepts (homonyms) occur depending on the disciplines. Supra-domain cultures shared by communities of language: Given a term-concept pair, terms (equivalents) vary depending on the target language.

From a knowledge engineering point of view, the domain-specific terminology encompasses term variants, which are realized as synonyms, homonyms and equivalents, which should be represented in the term bank. From a usability engineering point of view, the domain-specific terminology reveals the complexity of target-user expertise, which is shaped by communities of knowledge, practice and language.

Example of term variation around the concept VAT

Figure 3.2 shows the term variation motivated by the dimensions of domain cultures.

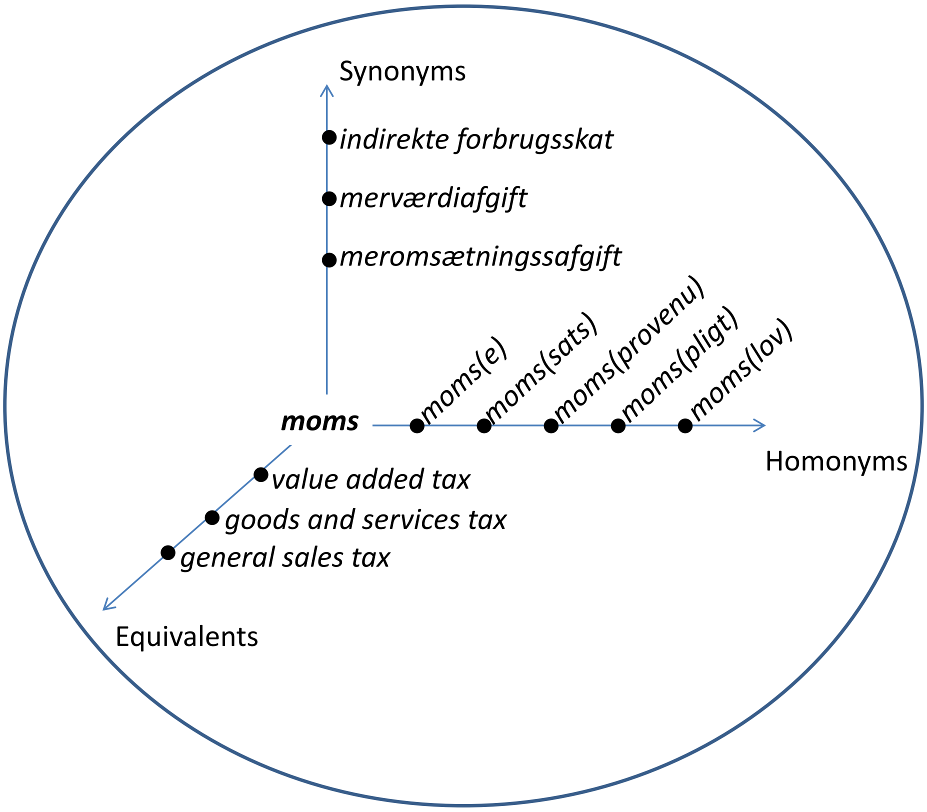


Figure 3.2: Examples of the domain-specific term variation. The term variation motivated by the dimensions of domain cultures around the concept (meaning) of value-added tax (*moms*).

Along the x-axis, communities of knowledge interpret moms into homomyms with different meanings: each community is concerned with ways to impose VAT, i.e. "to vat" (*momse*). Financial communities refer to moms as the rate (*momssats*) or the revenue (*momsprovenu*), while legal communities are likely to refer the liability (*momspligt*) or the law (*momslov*). Along the y-axis, communities of practice use synonyms to express the same meaning: Laypeople know the term moms, but increased sales tax (*meromsætningsafgift*), which was the original name abbreviated into moms, or tax on added value (*merværdiafgift*) may be used to semi-experts and indirect consumption tax (*indirekte forbrugsskat*) to experts. Along the z-axis, communities of language identify equivalents in the target language: The EU uses “value added tax” (VAT), while the OECD included “goods and services tax” (GST) and “general sales tax” (GST) in the world wide revenue statistics.

3.6 Knowledge acquisition from term banks

The primary purpose of a term bank is to give users access to terminological information from which they can acquire terminology and knowledge of specific domains. Besides ensuring performance of the knowledge tool, i.e. coverage (large quantity of the terminological data), consistency (high quality of the terminological data) and efficiency (high performance of the data base technology), we should develop a user interface with high usability ensuring effective user performance. Expertise of target users is complex due to domain cultures, but the domain-specific knowledge to be acquired is equally complex. Both sides are crucial to the performance of the eye-tracking experiment.

Research concerning human knowledge and cognition is massive and diverse. Let us begin by citing Solomon, Medin & Lynch (1999) asserting that “[c]oncepts are the building blocks of thought. How concepts are formed, used, and updated are, therefore, central questions in cognitive science.” Rusanen & Pöyhönen (2013) emphasize the processes underlying cognition, because “[b]riefly, in cognitive sciences concepts underlie complex cognitive states, such as thoughts or conceptions. Consequently, they are crucial to such psychological processes as categorization, inference, memory, learning, and decision-making [...].” Alexander (1992, p. 34) captures essential features of knowledge “domain knowledge encompasses declarative (knowing that), procedural (knowing how), and conditional (knowing when and where) knowledge,” based on which I choose to interpret performance of target users. I prefer Alexander's (1992) typology, but in experimental designs, we need to add a meta-level of cognition (or knowledge as I choose to label it) introduced by Kuhn (2000), see table 3.3.

Below I propose a separation of the cognitive processes underlying declarative knowledge, procedural knowledge, conditional knowledge and meta-knowledge. I acknowledge that there is potential overlap between the knowledge types, i.e. parallel processing (Balling, Helplund & Sjørup, 2014). Moreover, I acknowledge that the separation of general knowledge from domain-specific knowledge is controversial, as the parallel processing is likely to spill over in competent target users possessing sizable knowledge. In other words, the proposed separation of cognitive processes is merely a theoretical one, while in practice it is not possible to evaluate participants' performance of each of the types of knowledge.

I do not claim that the identified set of domain-specific knowledge types cover all mental activity, instead I suggest the four most crucial to the analysis of target-user performance, i.e. declarative knowledge (section 3.6.1), procedural knowledge (section 3.6.2), conditional knowledge (section 3.6.3) and meta-knowledge (section 3.6.4).

3.6.1 Declarative knowledge

Declarative knowledge is located in long-term memory (Rikers & Paas, 2005) and comprises the knowing of facts and definitions (Bonner & Walker, 1994). Methods to measure declarative knowledge include tasks where informants are asked to "declare", i.e. state terms, facts and definitions from their domain-specific knowledge (as opposed to explain conceptual relations or causes).

In the taxation domain, declarative knowledge is revealed when asking participants to mention one of the many excise duties in Denmark. In other words, participants with high declarative knowledge should be able to recall the terminological units, but not necessarily account for their meaning or the underlying conceptualizations.

It should be noted that there is an inherent challenge in the taxation domain, when measuring declarative knowledge, because nearly all target users are also taxpayers and will be exposed to the taxation terminology e.g. when filing their annual tax returns. Therefore, we should carefully choose the target terms of experimental design to alleviate the potential bias of target-user performance.

Seen from a term-bank view, declarative knowledge is the term-oriented content, see table 3.3.

Table 3.3: Domain-specific knowledge acquisition. The different knowledge types (1) and the skills (2) revealing each knowledge type and the domain specificity (3). Term-bank facilitation (4), eye-tracking experiment tasks (5), potential performance bias (6) of target users.

Knowledge type (1)	**Skill /Expertise (2)**	**Domain Specificity (3)**	**Term-bank facilitation (4)**	**Eye-tracking experiment (5)**	**Performance-bias (6)**
Declarative knowledge	State facts from the domain	Yes	Yes (term)	A-questions	Confused by the content
Procedural knowledge	Explain definitions from the domain	Yes	Yes (concept)	D-questions	Confused by the representation of content
Conditional knowledge	Apply problem-solving methods	Yes	No	N.a.	N.a.
		No		Experimental skills	
Meta-knowledge	Learn/Acquire knowledge	No	(Yes) (user adaption)	Trial-number effect	Confused or tired by the manipulations or material

3.6.2 Procedural knowledge

Procedural knowledge is located in the long-term memory (Rikers & Paas, 2005) and comprises the knowing of rules and procedures (Bonner & Walker, 1994) and knowledge of the actions necessary for problem solving in a particular domain, but not necessarily the expertise to conduct problem-solving (conditional knowledge). In my view, the long-term memory (Kalyuga & Sweller, 2004) is closely related to the notion of procedural knowledge as well as declarative knowledge, because long-term memory is defined as the ability of participants to categorize and store domain-specific knowledge into structures, from where it can be readily retrieved.

Methods to measure procedural knowledge include tasks where informants are asked to account for “procedures” crucial to the application domain i.e. the ability to explain connections or causes from their domain-specific knowledge (as opposed to merely state facts and definitions). It should be noted that informants may be able to repeat facts and definitions (declarative knowledge), without being able to understand underlying conceptualizations and explain connections and causes (procedural knowledge), whereas the opposite seems rather impossible.

In the taxation domain, procedural knowledge constitutes for instance the different principles for imposing direct and indirect taxation, which are the subject of a potential question in the experiment reflecting the task that a term bank may fulfill. In addition, procedural knowledge may be revealed by categorization tasks. The knowledge acquisition literature distinguishes between declarative and procedural knowledge, but in practice it is hard to draw a clear separation, as the two types of expertise spill over. Therefore, it is difficult to know if participants, who perform well (i.e. correct, deep and fast performance) in recalling taxation terms, which, at first hand, looks like an exercise of declarative knowledge (stating, not explaining, terms), might equally be reflecting procedural knowledge structures.

Seen from a term-bank view, procedural knowledge is the concept-oriented content, see table 3.3.

3.6.3 Conditional knowledge

Conditional knowledge or strategic knowledge is located in the working memory (Friege & Lind, 2006). Given their declarative and procedural (domain-specific) knowledge, their conditional knowledge will enable informants to solve tasks. Some researchers refer to the working memory as short-term memory (Kalyuga & Sweller, 2004). Graves (1996) describes the notion of expertise as both domain-specific as well as comprising "appropriate accompanying strategies." The skills to apply appropriate strategies may be categorized as conditional knowledge. In the absence of expertise, in so-called novices, "appropriate knowledge is missing" and only "general strategies" are available (Graves, 1996). Therefore, conditional knowledge is not necessarily domain-specific as opposed to the declarative and procedural knowledge types. However, in a term bank, it is not possible to facilitate target users with domain-specific problem-solving content as that would require e.g. models to calculate optimal indirect taxation.

It should be noted that, inference-making during experimental tasks reflects (general) conditional knowledge, which constitutes the way target users are able to reason about and solve a problem at hand, capable of actually producing an answer to the questions posed. In my view, inference-making, in line with meta-knowledge, should be seen as separate from declarative and procedural knowledge, which are domain-specific and located in long-term memory. In experimental terminology (see chapter 7), target users must instead apply general strategies (Graves, 1996) to complete the eye-tracking experiment.

3.6.4 Meta-knowledge

Meta-knowledge is, as the name indicates, the knowledge that an informant has about his own knowledge acquisition, in for instance an experimental setting. In other words, meta-knowledge is the (local) learning that may take place either about the experiment per se (general) or the term-bank content (domain-specific). It goes without saying that, as far as possible, a researcher should not disturb or confuse the meta-knowledge process. I should strive to ensure that briefing and instructions do not confuse participants, but encourage and make sure they are as comfortable as possible.

In my research, I analyse potential target users' performance by conducting an eye-tracking experiment. The eye-tracking equipment is located at an office (without natural light) at Copenhagen Business School, which in it-self proves challenging to informants, especially if they participate in eye-tracking experiments for the first time. I knew this might prove a disadvantage to at least half of the participants, namely staff-members from SKAT, and therefore I carried out information meetings to explain and illustrate what eye trackers are.

In experimental research, meta-cognition includes the mental activities the informant has about the intentions behind the experiment he is participating in, i.e. he offers the experimenter his attention instead of focusing on the task he has been asked to solve. Methods to measure meta-knowledge are retrospective interviews and other methods where the informant is offering his opinion about the experiment per se. In particular, when he ends up asking about what the point is, or when an informant suddenly refuses to do his best because the tasks seem pointless to him either because he is annoyed about what he anticipates will be interpreted by the experimenter as a poor performance, or because he is insecure about his performance, which might be kept secret to him due to the experimental design. The latter is very detrimental to data collection and should be avoided through thorough information and instruction of participants prior to the experiment. It is among other things necessary to ensure plausible scenarios, which informants are able to imagine and manage, as well as avoiding manipulations that disturbs a natural performance of informants (see table 3.3). The latter will reflect if informants learn as the experiment proceeds.

In research studying user behaviour, it is crucial that users are not aware of what is being observed, otherwise that might bias user behaviour. Therefore, we should take care not to reveal that during instructions. Afterwards in the debriefing, however, it is imperative to explain the whole experiment and repeat participant's option to withdraw their data. As an experimenter, it is hard to completely avoid that participants think about their performance and want to perform well, despite the fact that they might not have any clear notion of what “well” means. In order to avoid this parallel processing during the experiment, I chose to give users feedback on whether they answered correctly or not in each trial.

3.7 Conclusion

From a pragmatic point of view, I conclude that it is necessary to account for multiple domain cultures in accounting for the motivation for term variation.

I conclude that in the proposed framework, domain cultures reflect the motivations for term variation to overcome barriers related to practice (i.e. functional motivations), knowledge (cognitive motivations) or language (linguistic motivations). It follows directly from this framework that domain-specific expertise of target users will be multidimensional and complex, and expertise assessments would have to account for the competence in each of the three dimensions.

I conclude that target users sharing intra-domain culture share a practice and a corresponding communal lexicon, which is determined by their level of specialization, where the term variation (lexical signature) expresses knowledge employing different levels of abstraction, the so-called intra-domain registers.
Along the dimension of the intra-domain culture, expertise metrics may be linked to target users' occupation. Appropriate communication strategies must effectively overcome the barriers between the intra-domain cultures. In practice, we should

develop term banks that provide target users with terminology and knowledge to manage the different vocabularies belonging to the same conceptualization.

I conclude that target users sharing extra-domain culture share a scientific approach to a set of fundamental concepts, and term variation reflects distinct conceptualizations. Inside a community of knowledge, the interpretations are shared and stable, at least until new term formations are required to express new knowledge. Along the extra-domain culture, the target users' expertise must reflect the domain player's level of knowledge potential.

I conclude that target users sharing supra-domain culture share a language expressing the domain-specific terminology, but term variation (or translation strategies) is directional equivalence, because natural equivalence is often non-existing in legal domains. For instance, in the supra-domain taxation culture, expertise reflects the knowledge about other taxation systems at national, regional and global levels.

I conclude that the proposed graphic illustration displaying the dimensions of the cultures of domain-specific terminology (see figure 3.1) is useful for understanding the pragmatic view of disseminating knowledge to the potential target users. Disregarding the third axis (height), reduces the figure to the target-user matrix. Moreover, in the chosen taxation domain, we may use figure 3.1 to illustrate relevant examples (see figure 3.2).

I conclude that in term banks, the crucial knowledge acquisition processes facilitate target users with (domain-specific) declarative and procedural knowledge, while it is not aimed at (domain-specific) conditional knowledge. In the experiment, we should strive to avoid meta-knowledge confusing participants and reducing performance.

Regarding research theme, I conclude that the motivation for term variation is induced by the need to convey knowledge across different intra-, extra- and supra-domain cultures. The framework uncovers the potential barriers that target users strive to transcend in acquiring knowledge, which potentially furthers the development of a term bank by mapping target terminology (data), identifying target users (situations) and suggesting expertise metrics underlying user performance (needs). In other words, the multiple framework fully uncovers target users of the term bank.

Chapter 4: Experimental terminology

In this chapter, I outline the experiments aimed at examining the performance of target users of a term bank, constituting the third underlying research theme.

User-oriented research on interaction design and mixed methods are reviewed (section 4.1). Then I describe the chosen sample of participants, eye-tracking methods and tasks. In particular, the self-rating tasks, the representative tasks and the term bank tasks. I discuss the expected performance and validity issues of the experiments (sections 4.2-4.5). In addition, I present the results of the retrospective interview and the focus-group discussions (sections 4.6 and 4.7).

4.1 Introduction

First, my experimental design is guided by the limited-capacity assumption. Paavola, Lipponen & Hakkarainen (2004, p. 557) states that “knowledge is understood as a property or capacity of an individual mind,” which leads us to the limited-capacity presumption. It is assumed that the working memory is of limited capacity, which means that providing target users with redundant information may result in an overload of the limited-capacity working memory (Kalyuga & Sweller, 2004). Seufert (2003, p. 228) states that “[m]ultiple representations can serve many functions for learning. First, multiple representations may complement each other with regard to their content. For example, complex information is often distributed over multiple representations in order to avoid overloading a single representation. A second function is that multiple representations can complement each other with regard to their representational and computational efficiency, as different forms of representation may be differently useful for different purposes (Larkin & Simon, 1987).”

Moreover, it is the aim to inquire into expertise as means for user-adaption. Novice-expertise paradigm was introduced by chess studies e.g. asking participants to duplicate an arrangement of pieces (Fadde, 2009), which lead to my recalling task. Rikers & Paas (2005) propose expert characteristics based on their review of theories accounting for difference in performance between experts and novices. As Friege & Lind (2006, p. 442) state it, “neither too difficult for novices nor too easy to solve for experts.”

In research aimed at interaction design, there are generally speaking two strategies to follow: either we adapt the technology to the users, or we adapt (i.e. train) the users to the technology. It is a chicken and egg situation, where technologies must satisfy user needs, but where user needs are conversely shaped by technologies, so it is necessary to iterate between users and technologies. In practice, we should strive to minimize the discrepancy between the user and designer (Rubin & Chisnell, 2008), where the designer in so far as possible develops design with the target user in mind, but where the target user also needs to be instructed carefully to the functionalities of a new software tool, and actual use uncovers new usability problems, prompting another iteration of the interaction design process. The line of research in participatory design emphasizes this cooperation between designers and users throughout the design process (Simonsen & Robertson, 2013).

It should be noted that I avoid using the term “personas” (Nielsen, 2004). However, the motivation behind the development of personas is attractive: “to avoid the trap of designing for the ‘average’ user that doesn't actually exist, and instead to make sure that the system will work for somebody specific rather than no one in particular” (Usability First, 2015).

I propose a novel dual-entry mode representing domain-specific terminology in text and graphics to target users, whose expertise is depending on multiple domain cultures, which makes potential user adaption complex. In chapter 2, the combination of (term-oriented) lexicography and (concept-oriented) terminology renders possible the design of a dual-entry mode (prototype) consisting of a complementary article (text) and diagram (graphics) to be included in the user interface of the term bank. A dual-entry mode is novel and I expect that users will adapt to the prototype, instead of preferring the traditional single-mode entries of knowledge tools.

In chapter 3, I have outlined multiple domain cultures, which make expertise of target users complex. A common research strategy evaluating usability (Rubin & Chisnell, 2008) is to assess the performance of target users testing a prototype of a user interface, which is adapted according to predefined expertise levels, where users with high levels of expertise are assigned to a complex version of the prototype, while users with lower levels of expertise are assigned to a simple version. However, this strategy becomes difficult, because during the phase of developing expertise metrics, it is not possible to assign the user-adapted prototypes to the intended target users. Instead, I choose to postpone the user adaption and focus on the (direct and indirect) measures capturing expertise of target users (see chapters 5 and 6) and present the same (non-adapted) dual-entry mode prototype to all participants in the eye-tracking experiment (see chapter 7).

Numerous research areas concern visualisations. In particular, modern lexicographic research contributes to the knowledge on multimodal electronic dictionaries (see e.g. Lew, 2010) while different methods for visualising ontologies aim at providing users with overview (see e.g. Katifori, Halatsis, Lepouras, Vassilakis & Giannopoulou, 2007). I stress the duality of the proposed entry-modes containing both the diagram and the article, as opposed to testing the entry-modes separately. I expect the duality will facilitate target users in the acquisition of knowledge. Therefore, I focus on the static (non-adapted) images of the dual-entry mode to obtain a benchmark before introducing complex user adaption and dynamic functionalities. However, research on the presentation of meaning in dictionaries, (Lew, 2010) shows quite surprisingly that participants' access to animations had a negative effect on the ability to remember the meanings better.

Mixed methods

The primary motivation for combining the distinct quantitative and qualitative methods is target users' limited introspective access to the cognitive processes of knowledge acquisition (Britton, Stimson, Stennett & Gülgöz, 1998). To some extent, it would be a fruitful strategy to disregard target-user processing and ask potential users what they expect to be able to achieve by means of a term bank and what functionalities would be supportive in the user interface, especially, if the questioning is part of a survey conducted after the implementation of the term bank. The problem is that we are likely to encounter a discrepancy between what users actually need to achieve effective performance and what users think they need. The latter is constrained by users' limited knowledge about existing as well as future knowledge tools.

We may use the term “mixed” about the chosen research strategy, because the point of departure is the application of experimental methods to collect quantitative data, which are enriched by qualitative inquiries. This is sometimes referred to as triangulation. At the outset, triangulation was used to refer to the use of multiple qualitative and highly “interpretative” methods (Denzin, 2012). For a full review of the literature on the conceptualizations in mixed methods research and the image of a multidimensional crystal, see Denzin (2012).

There is not one fixed “set of methods that is appropriate; rather, the criteria for choosing methods include the following: What fits with the research question in this study” (Mertens, 2012). The combination of multiple methods will not ensure validity (i.e. ensure that we measure what we intended to), instead it adds depth, rigor and richness to the understanding of the object of study (Denzin, 2012). Strictly speaking only quantitative and qualitative data exist (Biesta, 2010 cited by Mertens 2012), but we use the terms quantitative and qualitative methods (or research), which result in the collection of quantitative and qualitative data. Usually, quantitative data are the result of measurements containing numerical or categorical data, while the qualitative data often is descriptive containing verbal data. However, hybrids exist, where quantitative data is generated by qualitative inquiry e.g. interviews. It should be noted that the context of study says not necessarily anything about the methods applied. Of course, it is often the case that field observations in a naturalistic context of participants generate qualitative data, while experimental designs necessarily must be de-contextualised to allow for the manipulation enabling us to isolate cause from effect.

Experimental design

My research aims at uncovering the behaviour of target users interacting with the proposed dual-entry modes to further the user interface design of a term bank. In particular, I assess the usability of the proposed dual-entry modes, among other things the performance in terms of correctness, speed and depth. The user behaviour will to some extent be introspective, which cannot be inquired into directly, but calls for methods where the cognitive processing can be inferred. I wish to emphasize the duality and include the graphics by means of concept diagrams in the interface together with written article entries into a dual-entry mode. In that case, we would not get far with asking users how they (think) they (would) perform, rather we must apply an experimental method to uncover users' performance and for that reason, I turn to cognitive science.

The established research methods of cognitive psychology are dominated by experimental designs with controlled laboratory tasks producing objective performance measures from which underlying cognitive processes are inferred (Ormerod & Ball, 2008). The advantage of designing and running experiments is that it makes it possible to manipulate key variables, thereby closely relating cause and effect by ruling out alternative causes. It is argued that mixing methods can sustain the relevance of highly controlled, de-contextualised (in-vitro) quantitative eye-tracking experiments with qualitative methods as it would emerge in its natural (in-vivo) context. See Dunbar and Blanchette (2001) for a summary of interesting research on scientific reasoning that is based on the triangulation naturalistic and laboratory methods in what they refer to as the “in vivo/in vitro” approach to cognition.

Experimental methods are well suited to isolate key effects on user behaviour. However, the strict manipulation of certain variables raises validity issues. First of all, we need to make sure that the experimental design allows us to “accurately attribute an observation to a specific cause rather than alternative causes” (Usability First, 2015). This is sometimes referred to as internal validity. This means that we should be very careful with the manipulation, as well as the hypotheses during the inferential statistics, including carefully separating the explanatory independent variables (causes) from the dependent variables subject to the potential effect, otherwise the experiment will not answer the research questions.

Visual attention

Remote eye trackers reduce the intrusiveness, and increase the ecological validity of the experiment (Loboda & Brusilovsky, 2010). I have chosen to apply eye-tracking technology in my experiments to investigate potential target-users' visual attention (Bundesen & Habekost, 2008) as evidence for the processing of the dual-entry-mode stimuli constituting low-fidelity design-artefacts of the term bank. Eye-tracking methods are widely used in dictionary research (see. e.g. Tono,

2011) as well as in user interface design research (see e.g. Goldberg, 2014). It has to my knowledge neither been used in the terminology field nor in term bank development.

It is a common problem in eye-tracking research that we face a discrepancy between participants on-screen eye movements (including their gaze) and their cognitive processing, because we cannot be entirely sure that what participants look at is also what is occupying their mind (Holmqvist, Nyström, Andersson, Dewhurst, Jarodzka & Van de Weijer, 2011), an issue spawning en entire research area of its own e.g. detecting events. Despite this discrepancy, I choose to apply the famous eye-mind hypothesis of Just & Carpenter (1980) implying that what the eyes are looking at is what the mind attends because there is presumably no appreciable lag between fixation and processing of target users. It is possible by means of mixed methods to collect additional auxiliary data to support and enrich the interpretation of the results indicated by the eye-tracking analysis.

4.2 Participants

The overall purpose of my research is to further the user interface design of a term bank, and I do among other things analyse the usability of the proposed dual-entry mode representing domain-specific terminology in text and graphics. The standardised definition of usability is the "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." (ISO, 1998). Therefore, it should be ensured that the results of the experiment generalize from the sample of users to the actual term-bank use, i.e. to the population of specified target users in the specified context of use where a specified goal is knowledge acquisition from interacting with the user interface of a term bank (product). I use the standardised usability approach as a starting point, but other aspects of term-bank use may be relevant. Term banks may also be developed in organizations, where collaboration is relevant, or targeted at consumer segments, where emotional user aspects are relevant (see figure 4.1).

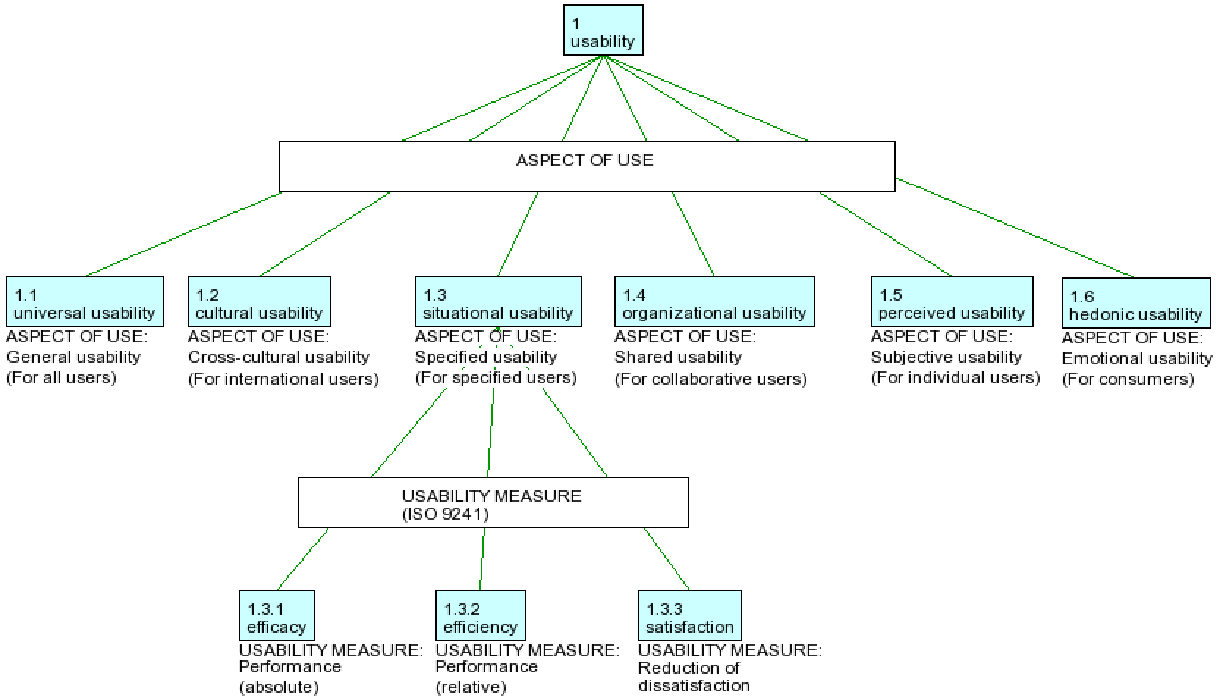


Figure 4.1: Usability ontology. The ontology displays Hertzum's (2010) six views on usability. The usability standard (ISO, 1998) is closely related to the situational usability (1.3), which emphasizes the specification of users and their goals and context. In comparison, organizational usability (1.4) emphasizes the shared and collaborative user situation. The four remaining views on usability comprise universal (1.1) and cultural (1.2) usabilities pertain to global target groups, while perceived (1.5) and hedonic (1.6) usabilities pertain to individual target users. The definition of usability is according to the ISO-standard the extent to which a product can be used by specified users to achieve specified goals with efficacy (1.3.1), efficiency (1.3.2) and satisfaction (1.3.3) in a specified context of use.

Professionals as target users

I have chosen the specified target users of the term bank to be professionals, which I define as the users who would access the term bank in connection with their work. Professionals are advanced members of the user groups outlined in the target user matrix (see table 3.1), i.e. they are expert members of different communities of knowledge or practice[3]. In my case study, professionals are e.g. employed as legal, financial or administrative personnel with advanced working tasks. Strictly speaking, expert is not synonymous with professional because

[3] I disregard professionals of language communities other than Danish (SL)

expertise is complex and professionals may exhibit different levels of expertise in each of the three proposed dimensions, i.e. a tax law expert may not exhibit high expertise in disseminating his knowledge to lay people or translating his domain to foreign languages.

There are a number of advantages of using professionals as target users. Firstly, professionals are most likely to represent the full scale of expertise in each of the proposed dimensions (see figure 3.1). This is not necessarily the case for non-professionals. Usually, user-oriented lexicographic research divides target users into experts, semi-experts and lay people with either communicative or cognitive needs (see e.g. Verlinde, Leroyer & Binon, 2009). Secondly, the object of study is the target-user processing of the proposed dual-entry modes which are novel to both professionals and non-professionals. Thirdly, the professionals constitute the target users with the best potential for promoting the term bank after launch, because of their widespread contacts with other user groups.

Validity

Experiments with target users raise at least two validity issues. Firstly, it is important to ensure an unbiased selection or sampling process. A selection bias is "any imperfection in the selection process that gets either the wrong types of users (people who aren't in your target audience) or a sample of users that is not random and therefore is biased toward one type of user rather than another (e.g. even though you want to target both novice and expert users, you only manage to recruit expert users, so your results will be skewed)" (Usability First, 2015).

In addition to the biases that the experimenter may risk inducing during the selection process, the experimenter always faces the obvious problem that participants must volunteer, but not all members of the potential target users are open for recruitment. "The most common example is a volunteer bias, since people who volunteer to be in a study may not be representative of your overall target audience; they may, for instance, have more free time, have a greater interest in technology, and be more outgoing than those who wouldn't volunteer" (Usability First, 2015). Apart from issues in the process of sampling participants among target users, we need to make sure that we recruit sufficiently many participants to ensure the statistical validity in the analysis, i.e. that results can be relied upon, because they are not the result of random error.

Forty volunteers constitute the sample of participants for my experiments, and they were recruited by pursuing a sampling strategy, which concerned key expertise variables (see section 5.2). The "degree of statistical significance of a result depends upon the number of sampled data points (e.g. the number of users in a test), the size of the effect, and the amount of variation between measurements. Thus, statistical validity can be improved with increased samples, by reducing noise (and thus variation), and by creating as large an effect as possible" (Usability First, 2015). Prior to the actual experiments, five volunteers (one from SKAT and four from CBS) ran pilot versions, which did not cause any major adjustments to the design.

The chosen terminology experiments comprise self-rating tasks (see section 4.3) directly assessing expertise of participants and representative tasks (see section 4.4) indirectly assessing expertise. In addition, the term bank tasks are designed to capture target users' performance using the dual-entry-mode prototypes (section 4.5), while the retrospective interview and focus-group discussion concerned target users' own perceptions and preferences (sections 4.6 and 4.7). The experimental material was presented in Danish, but is translated in English in appendices A-G.

4.3 Self-rating tasks

The choice of self-rating was primarily motivated and inspired by the mental workload assessments developed by Hart & Staveland (1988), who used multiple sub-scales to reflect the workload that users experienced. In my view, when using information tools, the sub-scale on mental demand, i.e. the mental activity necessary for solving a task (Hart, 2006), seemed appropriate for the assessment of domain-specific expertise.

In the background questionnaire (see table A.2 in appendix A), participants are asked to self-rate their domain knowledge on a scale, i.e. an 80 mm line with "very little" at one end and "very high" at the other, and with an indicator at the middle (40 mm), see figure 4.2.

Figure 4.2: Self-rating scale of expertise. Participants were asked how big their knowledge is in the taxation domain and to indicate the level on the line ranging from "very small" (*meget lidt*) to "very large" (*rigtig meget*). The self-rating task was part of the background questionnaire (see table A.2 in appendix A).

The exact wording of the question was: "How large is your knowledge in the taxation domain? Please put a mark on the line below." Compared to the workload index, the question is slightly rephrased from the qualitative notion underlying "how complex" to the quantitative notion "how large" to avoid reversing and biasing especially the experts' self-rating:

If the complexity notion had been maintained, we would expect experts to rate themselves low (i.e. they find the taxation domain non-complex), while the non-experts would rate themselves high[4]. Moreover, it is equally possible that some experts rate themselves high (i.e. they find the taxation domain complex), which reflects that some areas are actually complex (e.g. taxation of real property) and this ambiguity had to be alleviated. The proposed rephrasing "how large" was meant to immediately prompt an intuitive and supposedly precise answer. As we shall see below, participants' understanding of the question "quantifying" their expertise, was not without ambiguity either (see section 5.3). In comparison, the self-rating tasks concerning English language and information searching were phrased as "how good are your skills," which turned out less biased (see section 5.4).

The advantage of self-rating is the subjective response, directly assessing a variable, which in theory was expected to be significant in explaining differences in performance levels. The self-rating was located at the beginning of the questionnaire, before any introduction to terminology work or representative-task

[4] In practice, it is not a problem with reversed variables per se, but it may challenge the interpretations unnecessarily.

exposure from the recalling and categorizing tasks (also part of the background questionnaire) to avoid any bias of participants' intuition, when rating their knowledge levels. The rating scale was intended to make participants level their knowledge continuously and relatively, but it is not meaningful to directly add numbers to the scale. In particular, the scale was deliberately not numerically designed e.g. beginning with absent knowledge (0%) to “all” knowledge (100%). In addition, I avoided the simplistic categorization of assigning target users into experts, semi-experts and laymen onto the scale, as unnecessary grouping reduces the statistical power (see section 5.3.2).

Apart from the domain-specific expertise, participants are also asked to self-rate their reference skills (see section 5.4) and their English language skills. The means, standard deviations and ranges of the answers are summarized in table 5.1. The self-rated domain-expertise variable is being validated by the proposed set of (explanatory) expertise variables in a regression approach and included as explanatory variables for participants' performance in the dual-entry mode experiments (see table 7.1).

4.4 Representative tasks

Participants performed the proposed set of four representative tasks aimed at reflecting the complex domain-specific expertise: The choice of representative tasks is highly motivated by terminology theory discussed by Cabré (2003), where the fundamental criteria require that terms are lexical units designating underlying concepts occurring in the specialized discourse (see section 2.2):

(1) The (open-ended) recalling task is intended to reflect participants' ability to produce a list of lexical units (words or phrases), which count as domain-specific terms i.e. designates an underlying domain-specific concept (see section 6.2).

(2) The categorizing task is intended to reflect participants' ability to recognize and define the underlying concepts designated by a given list of terms from a given list of pseudo-terms, which do not exist, i.e. designate any concepts (see section 6.3).

(3) The reading task is intended to reflect participants' ability to comprehend a piece of written specialized discourse in the form of a chosen, authentic text containing domain-specific terms (also part of the scenario of the eye-tracking experiment) (see section 6.4).

(4) The card-sorting task is intended to reflect participants' ability to represent (draw) conceptual knowledge structures of terms from the dual-entry modes in the diagram format (see section 6.5).

The representative tasks are evaluated by appropriate performance indicators, which were motivated by the expert characteristics discussed by Rikers & Paas (2005, p. 145) ascertaining that "experts are faster than novices at performing skills; experts perform their tasks (almost) error free; experts have superior short-term memory and long-term memory; and experts' problem representation is deeper, more principled, than that of novices, who tend to build superficial representations of a problem. First, it is assumed that the proposed representative tasks trigger short- and long-term memory (discussed in sections 4.4.1-4.4.4), which will be superior for experts. Second, I infer that the relevant performance indicators are correctness, speed and depth (see table 4.1).

Table 4.1: Performance indicators and experimental tasks. Recalling is primarily evaluated by means of depth. Categorizing is primarily evaluated by means of correctness. Reading is primarily evaluated by means of speed. Structuring is primarily evaluated by means of depth. "N.a." reflects that the data were not collected in the experiment. Parentheses reflect that the evaluation of the data was not meaningful and the task has to be improved.

Task	Correctness	Speed	Depth
Recalling	(Term list)	N.a.	Frequency
Categorizing	Existence	N.a.	(Definitions)
Reading	N.a.	Time	(AOIMost)
Structuring	(Reference map)	N.a.	Interconnectedness
Dual-entry mode	Correctness model (section 7.3)	Response-time model (section 7.4)	Diagram-fixation model (section 7.5)

The experimental design of tasks should allow for the evaluation of all relevant expert performance indicators. In practice, we are facing different constraints limiting the evaluation of performance.

4.4.1 Recalling

In recalling taxation terms, participants were asked to list the first ten terms from the taxation domain that came to their mind. The recalling of terminology is in practice not necessarily a matter of retrieving stored terminology. It is equally possible that terms are produced from a few key terms by means of association. The task of recalling taxation terms is located directly after the introduction to concept clarification in the background questionnaire, but before the subsequent categorization task to avoid any bias from exposure to useful taxation terminology prior to this test.

The test was initially included in the part of the experiment conducted in front of the eye tracker, which allowed for recording of response time. However, the pilot studies revealed a need to minimize the time in front of the eye tracker and to keep this test in the background questionnaire, hence discarding the response times, which were of secondary interest. It was also very important that the overall completion time of the background questionnaire was kept below 30 minutes, and the pilots suggested that asking for maximum ten terms was appropriate.

In other words, the recalling task is asking participants to state domain-specific terminology, which primarily reflects the declarative knowledge stored in the long-term memory. Expert performance is indicated by a frequency analysis of the produced term lists, primarily reflecting depth (see table 4.1). The analysis of the ten-terms recalling task is presented and discussed in section 6.2. The exact design of the task is presented and translated into English in table B.2 in appendix B.

4.4.2 Categorizing

The categorization task is inspired by lexical decision (typically focusing on response time and error rates for various conditions for words). Participants were presented with both existing terms and non-existing (so-called) pseudo-terms, which are semantically plausible expressions. A list of 15 randomly ordered

Danish terms and pseudo-terms were produced: Ten terms are members of the taxation discourse and five expressions plausible members of the taxation discourse (pseudo-terms), but constructed for this experiment (see table 6.5).

The selection criteria for the existing terms on the list are ontologically and morphologically motivated: Hence, the terms cover various positions in the terminological ontology: At the central government level: “tax receipts” (*skatteindtægt*) (see figure II.2 in appendix II), “tax revenue” (*skatteprovenu*), and “fiscal tax” (*fiskal skat*). Direct taxes (see figure II.3) collected from persons: “personal income tax” (*personlig indkomstskat*) (see figure II.4) and “middle-bracket tax” (*mellemskat*) (see figure II.5); from companies: “corporation tax” (*virksomhedsskat*) (see figure II.4) or from both: “land tax” (*ejendomsskat*) (see figure II.6). Indirect taxes “green tax” (*grøn afgift*), “energy duty” (*energiafgift*) (see figure II.12), and “consumption-limiting duty” (*forbrugsbegrænsende afgift*).

The five constructed pseudo-terms were semantically plausible (and different from non-words in lexical decision which tend to be orthographically and phonotactically nonsense strings), i.e. they seem to fit into the existing terminology. “C-tax” (*C-skat*) is supposed to get participants to think of the existing “A-tax” (*A-skat*) and “B-tax” (*B-skat*); “provisional deduction” (*forskudsfradrag*) seems likely as there is a “preliminary tax” (*forskudsskat*) and “preliminary assessment” (*forskudsopgørelse*); “coffee VAT” (*kaffemoms*) might resemble the existing “coffee duty” (*kaffeafgift*) or “bus VAT” (*busmoms*); “value-subtracted tax” (*mindreværdiafgift*) could seem the opposite of “value-added tax” (*merværdiafgift*); and “wind duty” (*vindafgift*) resembles the existing “water duty” (*vandafgift*) or “wine duty” (*vinafgift*). In addition, it was important to ensure that the most frequent domain-specific stems were present including “tax” (*skat*), “deduction” (*fradrag*), “VAT” (*moms*) and “duty” (*afgift*).

The pilot studies confirmed that it was necessary to include non-existing pseudo-terms as well as an additional task of defining terms to get participants to slow down and sincerely consider their answer, as the experimental list of terms and pseudo-terms are prone to being quickly categorized as existing without much careful consideration, and I therefore include a definition task and discarded response time (including feedback on speed).

To sum up, the categorization task is asking participants to explain domain-specific terminology, which requires procedural knowledge stored in the long-term memory. Expert performance is indicated primarily by correctness of the categorized terms (see table 4.1). The analysis of the categorization task of terms and pseudo-terms is presented and discussed in section 6.3. The exact design of the categorization task is presented in table B.3 in appendix B.

4.4.3 Reading

In the reading task, participants were asked to read an authentic, specialized text from the taxation domain. I chose a text published by Statistics Denmark (Statistics Denmark, 2012, p. 13), which also provided terminological information for some of the article entry modes (see figures D.3, D.6, D.8 and D.9 in appendix D) of the dual-entry mode experiment. The text includes terms from the taxation domain, e.g. “tax revenue” (*skatteprovenu*), “consumption-limiting duties” (*forbrugsbegrænsende afgifter*) and “income taxation” (*indkomstbeskatningen*). But less specific terms from the related public finance domain, e.g. “public spending” (*offentlige udgifter*), “gross domestic product” (*bruttonational-produktet*) and “market prices” (*markedspriser*) are also present. Finally, words or non-specific terms, e.g. “level” (*niveau*), “need” (*behov*) and “years” (*årrække*) are present in the text (see figure C.3 in appendix C).

An authentic text is used to ensure the statistical validity, i.e. that results generalize to the specialized texts. The text does not allow for a controlled comparison of participants' processing of domain-specific terms compared to other terms and words in the text for a number of reasons: Firstly, processing time of domain-specific terms does not necessarily follow the pattern predicted by reading research on generalized comparable words, since the (typically) low frequent[5], compound terms, which in theory are highly unpredictable may not necessarily imply difficulty and long processing times, which is the case with e.g. “consumption-limiting duties” (*forbrugsbegrænsende afgifter*). Secondly, it was not in practice feasible to identify comparable words in the text, where no obvious candidates were to be found (the words are typically not long enough, or not located in a comparable place in the sentence). Consequently, I focus on two

[5] Domain-specific terms often exhibit low frequency in a general-language corpus.

sections of the text, one is the most complex/deep (AOIMost) and the other is the least complex/deep (AOILeast) (see figure C.4 in appendix C).

To sum up, the reading task is asking participants to activate their short-term working memory. Expert performance is indicated by primarily reading speed (see table 4.1). The analysis of the reading task on a specialized text is presented and discussed in section 6.4. The exact design of the stimuli displayed on-screen is shown in figure C.3 in appendix C.

4.4.4 Card-sorting

At the workshop, which was held a few months after the eye-tracking experiments, nine participants (eight from the experiments and one not taking part in the prior experiments) did the card-sorting exercise. Card-sorting is "a technique for uncovering the hierarchical structure in a set of concepts by having users group items written on a set of cards, often used, for instance, to work out the organization of a website," with the option of pre-defined categories (Usability First, 2015).

The card-sorting exercise constituted 32 terms, 16 belonging to direct taxation and 16 belonging to indirect taxation. Eight terms of the 32 terms were target terms in the dual-entry mode experiment, which means that participants had encountered the terms 4-6 months earlier, when completing the experiment. I did not allow any warm-up as participants should have familiarized themselves with conceptual diagrams during the experiment and more importantly, I was interested in seeing participants' intuitive responses without any unnecessary conceptual exposure. In particular, I asked participants to use all the terms on the list, not to include new terms, and to draw one or several diagrams containing the terms and the relations between them. Unfortunately, I was not able to recruit all the participants from the dual-entry mode experiments to conduct this card-sorting exercise. The low number of participants produces a low number of observations, which proves problematic to the inferential statistics.

I expected that an expert would be able to reproduce a structure of the terminology reflecting the conceptualization of his domain. The concept mapping is not novel in terminology. Grabar, Hamon & Bodenreider (2012) refers to a study on the categorization of terms according to the expertise level of participants. It is shown that “a wide range of views is held by participants about the nature and structure of things within a domain. For instance, when asked to categorize terms, the participants demonstrate a diversity of views.

It was not possible to develop a meaningful independent variable, as the concept maps were highly subjective and differed very much from the reference maps, requiring qualitative interpretations. I limit the analysis to descriptive evaluations of the produced concept maps.

To sum up, the categorizing task would trigger both declarative and procedural knowledge stored in the long-term memory. Expert performance is primarily indicated by depth of the concept maps (see table 4.1). The analysis of the card-sorting task on a given list of terms is presented and discussed in section 6.5. The exact design of the card-sorting task is presented in table F.1 and the list of terms is outlined in figure F.1 in appendix F.

4.5 Term-bank tasks

The choice of term-bank tasks was primarily motivated by user-oriented research in interface design, where the usability testing of low-fidelity prototypes further the subsequent and iterative interface-design process.

The key thing in experimental design, which necessarily is de-contextualised due to the need to manipulate key variables, is to trigger a natural behaviour of participants. This requires that the design of the experiment provides users with a scenario that is easy to understand and matches the context of an actual term bank, without causing confusion that will distort participants' natural behaviour and performance. In the beginning of the eye-tracking experiment, I presented participants with a scenario asking them to imagine a plausible context of use. I chose a scenario outlining a work situation on a newspaper covering the European economy (see figure C.2 in appendix C), and the task was to write about the Danish taxation system (see figure D.1 in appendix D). The advantage of this scenario is that it seems plausible to the target users without exactly matching any of the participants’ job functions directly.

The advantage of the experimental approach is that we are able to design provisional dual-entry modes of the term bank, which include the concept diagrams and manipulate the questioning to ensure that participants search both diagrams and articles. Below I present and discuss the dual-entry modes (section 4.5.1), the chosen question types and target terms (section 4.5.2) as well as the procedure of the eye-tracking experiment (section 4.5.3).

4.5.1 Dual-entry mode (stimulus pair)

The purpose of the dual-entry-mode prototype is to test the usability and to determine whether the dual-entry modes should be part of the user interface of a term bank.

The design of the textual and graphical entry modes resemble an existing graphical user interface (GUI) of a terminology and knowledge management tool, *i-Term*, which supports the manual construction of terminological ontologies (graphics), but also enables the display of the entered terminological information in a textual format. In practice, the textual and graphical artefacts were produced in PowerPoint to be compatible with the experimental psychology software tool, *E-prime*. The proposed dual-entry modes were displayed concurrently (see figure 4.3) complementary access points (in texts and graphics) of a term bank interface.

Hvilken type skat eller afgift er energiskat?

1. **Miljøafgift**
2. **Energiafgift**
3. **Punktafgift**

Dansk: Kilde til term:	**energiskat** Skatter og afgifter 2012 (Danmarks Statistik)
Definition: Kilde:	punktafgift som pålægges varer og tjenester med det formål at begrænse miljøskadeligt energiforbrug DanTermBank
Kommentar: Kilde:	Den største gruppe blandt punkt-afgifterne er energiskatterne, der i 2011 tegnede sig for 44 pct. af samtlige punktafgifter. Skatter og afgifter 2012 (Danmarks Statistik)
English: Source for term:	**energy tax** Taxation trends in the EU 2012 (Eurostat)
Definition: Source:	excise duty imposed on energy consumption to reduce emissions DanTermBank

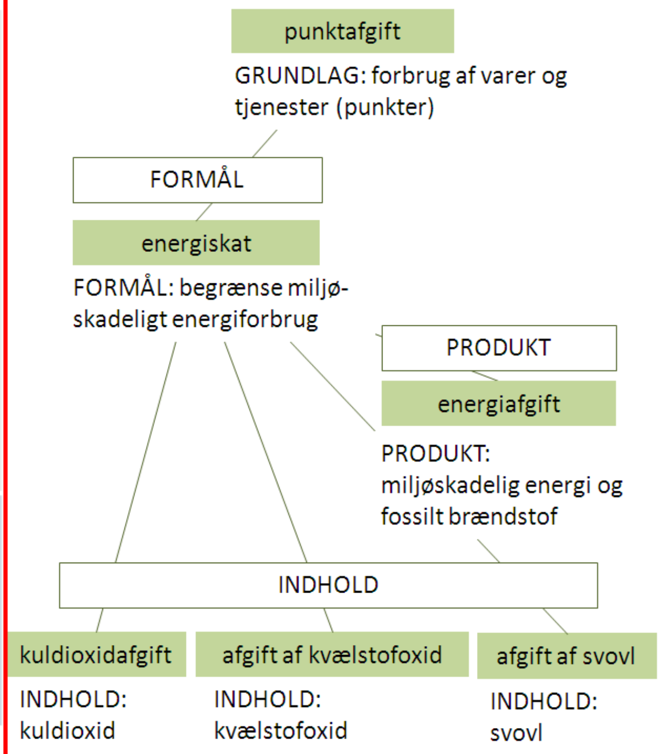


Figure 4.3: Example of dual-entry mode. The dual-entry mode of block 1 “energy tax” (*energiskat*), question type DA1 (type) and the diagram displayed to the right. This instance would imply that the corresonding question type (DA2) would have the diagram displayed to the left. For translation into English, see figure D.3 in the appendix. For the dual-entry modes of the remaining blocks 2-8, see figures D.4-D.10.

Target terms of the experiment were supposed to constitute potential entry words that target users look up in the term bank in either entry mode (by means of text or graphics). Despite the fact that terms constitute lexical units, they are not (like words) necessarily sensitive to complexity indicators e.g. frequency of general language corpus, length or predictability (see chapter 6). Therefore , I based the choice of target terms on the position of the chosen terms in the terminological ontology, i.e. I chose eight Danish taxation terms (organized in so-called blocks), four belonging to indirect taxation: “energy tax” (*energiskat*), “motor vehicles tax” (*afgift af motorkøretøj*), “green tax” (*grøn afgift*) and “excise duty” (*punktafgift*), and four belonging to direct taxation: “middle-bracket tax” (*mellemskat*), “land tax” (*ejendomsskat*), “personal income tax” (*personskat*) and “direct tax” (*direkte skat*) (see figures D.3-D.10 in appendix D).

To ensure comparability across the eight blocks and compatibility with the eye-tracking system, the textual article-entry and the graphical diagram-entry comply with entry-mode templates allowing for very little difference in the layout: The textual entry mode is displayed as a bilingual written article in tabular format with two columns and ten rows representing the terminological information (term, definition, equivalent, and comment including sources of both the Danish SL and the English TL), see the left panel of figure 4.3. The graphical entry mode displays information in a conceptual diagram including 5-7 concepts (nodes) structured in three levels of subdivision criteria. I only apply type relations resulting in a taxonomic structure of the concept diagrams, see the right panel of figure 4.3. The dual-entry modes are displayed concurrently in pairs of text and graph inside the stimulus space of the screen below the multiple choice questions (see figure 4.3).

4.5.2 Question types, format and feedback

Question types were designed and manipulated carefully to ensure that participants consult both the concept diagrams and articles, otherwise we would have run the risk that participants searched only their most preferred entry mode, or their most preferred side of the screen. Research on multimedia in learning environments have shown that users often are not able to use the multiple representations (in text, picture of graphics) effectively, especially users with low prior knowledge have problems with understanding multiple representations, and prefer to concentrate on the most familiar representation (Seufert, 2003). The results are confirmed by dictionary researchers that animations distract participants (Lew, 2010).

In total the experiment contained 48 questions. For each of the eight target terms (blocks), participants got six multiple-choice question types with three available answers for each (see table 4.2).

Table 4.2: Question types belonging to each block. The six question types of block 1 “energy tax” (*energiskat*) and the available answers (multiple-choice format) translated into English (see table D.1 in appendix D for the Danish versions). For questions belonging to the remaining blocks 2-8, see tables D.2-D.8. The answers available to question type A1 are necessarily in Danish (not translated).

Question type	Term-bank content	Question	Available answer	Correct answer
Diagram (D1)	Sub-ordinates	How many types of energy taxes exist?	1: Four 2: Six 3: Eight	No. 1
Diagram (D2)	Sub-division criteria	What separates carbondioxide tax from duty on nitrogen oxides?	1: Purpose 2: Content 3: Taxpayer	No. 2
Article (A1)	Equivalents	What can ’energy tax’ be translated into in Danish?	1: *energiafgift* 2: *energiskat* 3: *energitakst*	No. 2
Article (A2)	Comments	Energy taxes constituted 44 per cent of excise duties in 2011 according to whom?	1: OECD 2: Eurostat 3: Statistics Denmark	No. 3
Diagram-Article (DA1)	Super-ordinate (Definition)	What type of tax or duty is energy tax?	1: Environmental duty 2: Energy duty 3: Excise duty	No. 3
Diagram-Article (DA2)	Attributes (Definition)	What is the purpose of energy tax?	1: Limiting environmentally damaging energy consumption 2: Limiting environmentally damaging consumption 3: Limiting environmental Damage	No. 1

The six question-types fall into the following categories: The first diagram-based question (denoted “D1”) concerns sub-ordinates and the second diagram-based question (denoted “D2”) concerns sub-division criteria. The first article-based question (denoted “A1”) concerns equivalence and the second article-based question (denoted “A2”) concerns comments. The first diagram-and-article-based question (denoted “DA1”) concerns super-ordinate and the second diagram-and-article-based question (denoted “DA2”) concerns attributes (see table 4.2). The question types are motivated by the knowledge typically sought in and acquired from term banks e.g. equivalents in the target language, comments describing key features of target terms or definitions (i.e. question types A1, A2, DA1 and DA2) combined with information to be found in the concept diagrams (i.e. question types D1 and D2). As with the available (multiple-choice) answers, the question types were supposed to resemble plausible term-bank use reflecting realistic domain-specific information needs of potential target users.

The so-called diagram-based questions (denoted "D") can only be answered by consulting the concept-diagram stimulus, while the so-called article-based questions (denoted "A") can only be answered by consulting the concept-article stimulus, and finally the so-called diagram-article-based question (denoted "DA") can be answered by consulting either stimulus (see figure 4.3 and table 4.2). In other words, the answer to a question will be available in one of the three offered answers, but can also be found from viewing the stimulus-pair consulting information in only one of the entry modes or in either, though this might occasionally be questioned directly by the participants, one (no. 29) states it clearly: "Is the answer appearing from the screen?" (*Skulle svaret fremgå af skærmen?*).

Of course, it is also constantly an option for participants not to consult the concept diagram or article at all and answer immediately by choosing among the available answers, as one of the participants (no. 9) verbally expresses: "Am I supposed to read? I answered without looking!" (*Skal jeg læse? Jeg svarede uden at kigge!*). However, this was rarely the case, as one participant (no. 21) states quite clearly that even knowing the answer makes one doubtful and prone to read: "I know that but you have to read!" (*Det ved jeg godt, men man er nødt til at læse!*).

Multiple-choice question format

As mentioned, the questions are in a multiple-choice format with three possible answers, placed in a horizontal field above the stimulus-pairs, separated by thick red lines (see figure 4.3). The multiple-choice question format was chosen to keep the answering process (and the correctness evaluation of answers) as simple and systematic as possible, also avoiding time-consuming qualitative coding. Hence, instead of having participants formulate and type in (long, error-prone) answers in a textbox, they were asked to pick one of the numbers “1”, “2” or “3” to represent the correct answer. There was one and only one correct answer to each question. On the one hand, this format eliminated blanks, avoiding participants giving up. On the other hand, it cannot be rejected that (wild) guesses became a chosen strategy in cases of confusion or not-knowing-the-answer, as one participant (no. 6) puts it: “I am completely lost here. It is pure guesswork!” (*Her er jeg helt lost. Det bliver et rent gæt!*).

However, there is no reason to assume this drawback being more prevalent in the multiple-choice format than it would have been in an alternative (open-ended) self-formulating-answers format. It should be noted that multiple-choice questions are not only aimed at providing correct answers, they also provide wrong answers. Thus, participants are forced to carefully consider plausible alternatives, which may closely resemble the concept-clarification or knowledge-acquisition need driving the search of a term-bank target user. This means that this question format will imply that participants are expected to spend a considerable length of time processing the question, as one participant (no. 2) so precisely puts it: “Let me first understand what you are asking about!” (*Jeg skal lige forstå, hvad det er du spørger om!*).

Finally, evaluating the correctness of multiple-choice questions is a much simpler task than evaluating (and coding) answers formulated in prose, where degrees of correctness are likely to appear, as was the case in both the recalling and categorizing tasks (see chapter 6). Moreover, the multiple-choice question format allows feedback to parcipants.

Question feedback

When participants have answered the question, the system tells them, whether they answered correctly or not. This feedback feature was much appreciated during the pilot studies, providing an instructional factor at the beginning of the experiment, where participants familiarized themselves with the experimental format, and a motivational factor towards the end of the experiment, where participants showed signs of fatigue. Moreover, it turned out that feedback which surprised the participants triggered verbal responses, especially in the case of disagreement, which could then be followed up in the retrospection, after the completion of the eye-tracking experiment.

The retrospective questions asking participants to evaluate their performance would not have been possible without an indication of errors. However, the feedback had two (minor) downsides: First, the answering time will be prolonged by previous wrong answer (triggering insecurity as one participant (no. 3) precisely states it: “It takes time to do it right!” (*Det tager tid at gøre det rigtigt*!) and shortened by a previous right answer (triggering confidence), which (in principle) biases a naturalistic target-user behaviour, where the feedback is unavailable. Second, in the field of usability, it is stressed that user behaviour is not (only) a matter of answering right and wrong, rather the system should be capable of encompassing different kinds of users, at least in the case of adapting the information technology to target users.

Concept-diagram focus

All participants are answering the multiple-choice questions with the relevant dual-entry modes displayed. It depends on the question type, whether a particular entry mode contains the answer. This design allows for the analysis of each main type of question (denoted “D”, “A” and “DA”) as well as participants' fixations on each displayed entry mode (concept diagram and concept article) separately. But more importantly, I ensure that participants need to search the diagram (the entry mode least familiar to them and for that reason presumably less preferred), which is the case when the answer is only displayed in the diagram (D-question). For comparison, I include a situation, where the answer is displayed both in the diagram and article (DA-question) as well as a situation, where the answer can only be found in the article (A-question).

The experimental design allows me to emphasize the duality of the entry modes, i.e. whether participants are able to manage the dual display or whether the diagram is confusing participants and reducing performance and should be avoided in the user interface. I choose not to construct a control experiment, where participants are completing a simplified version of the experiment without any entry modes (just questions) or with only one entry mode displayed at a time, because those performance levels are not useful for my analysis. Overall, the development of performance across trials (learning effects) is more important than alternative performance in comparable single-entry or no-entry questions. In the view of user interface design, learning effects of the dual-entry modes may be interpreted as the case, where users adapt to the low-fidelity prototype. In particular, I investigate expertise effects on target users' performance, which may be interpreted as the basis for user-adaptive visualizations (Loboda & Brusilovsky, 2010), i.e. adapt to users.

4.5.3 Experimental procedure

Prior to the dual-entry-mode experiment, participants answered a background questionnaire and completed the three self-rating tasks (i.e. expertise, reference skills and English language skills), as well as two of the four representative tasks (i.e. recalling and categorizing).

Eye-tracking experiment

Participants did the eye-tracking experiment in an eye-tracking lab. The reading task (see figure C.3 in appendix C) was part of the instruction and scenario (see figure C.2 in appendix C), after which the dual-entry-mode experiement began (see figure D.1 of appendix D).

A remote *SensoMotoric Instrument (SMI)* eye tracker, which supports gaze sampling rates of 50 Hz, is used for the recordings of participants' on-screen eye-movements during the experiment. This means that the eye tracker records gaze directions 50 times per second, which in the eye-tracking world is a rather low frequency. In this experiment, however, the primary areas-of-interest (AOI) are the entry-modes (or stimulus-pairs), which constitute a large part of the screen (see figure 4.3). AOIs are spatial, delimited areas (in this case nicely shaped like boxes), applied onto the stimuli using *BeGaze*, which then reports when and how long the eye gazes on the AOI. The large AOIs combined with fixation thresholds above 200 ms, the low-frequent eye tracker producing low gaze sampling rates should be sufficient, see figure 4.4.

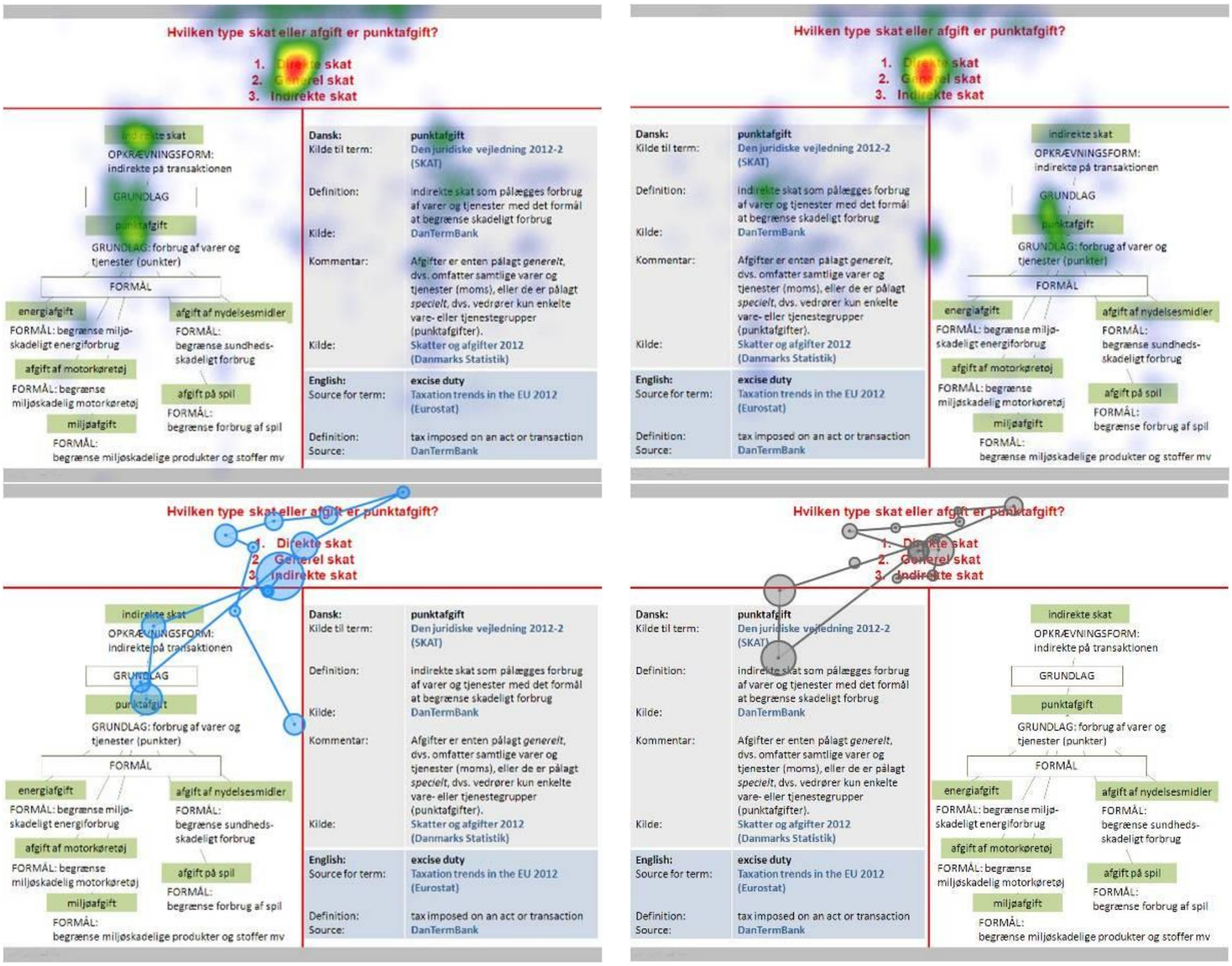


Figure 4.4: Visualization of eye movements. The dual-entry mode concerning "excise duty" (*punktafgift*) (Block 4) and question concerning definition (DA1). Top left panel: Heat map with the diagram to the left (18 participants). Top right panel: Heat map with the diagram to the right (19 participants) Bottom left panel: Scan path of the fastest participant with the diagram to left (4.2 seconds, trial number 23) Bottom right panel: Scan path of the fastest participant with the diagram to the right (4.4 seconds, trial number 33).

The experiment is built in the psychology software *E-prime*, which facilitates randomization, controls the eye tracker, records user responses, and informs participants whether they answered correctly or not (feedback). This design means that if the eye tracker suddenly stops recording (for a longer or shorter time), the answers and response times of participants are still collected from all participants and used for analyses (see sections 7.3 and 7.4). Moreover, *E-prime* controls the two manual 13-point calibrations. The first one placed after the instructions and the second one placed half-way through the 48 questions. In particular, the latter calibration made it possible to maintain or even improve the quality of the eye-tracking data by changing participants' distance to the screen or re-centering their position in front of the screen, before the final part of the experiment continued. In experiments with remote eye trackers, it is crucial to calibrate the system, as participants are free to shift position, which is not the case for tower-mounted eye trackers, where participants' heads are in a fixed position, or the (mobile) head-mounted eye trackers, where the camera is fixed to the head in a helmet or a pair of glasses. Remote eye-tracking system was the most suitable for this type of experiment resembling a relatively naturalistic user situation of a term-bank target user focusing on relatively large AOIs.

Randomization

All participants answered the same 48 questions but in different random orders. Randomization is conducted by *E-prime* during each session, at three levels: side, question and block.

Firstly, the display-side of the diagram of the first question type, i.e. D1, A1 or DA1, is randomly assigned to the right or left, leaving the opposite display-side to the second corresponding question type, D2, A2 or DA2, respectively. This display-side randomization is included to avoid any biases from the preferred viewing behaviour for reading in Danish going from left to right, top to bottom.

Secondly, the six questions inside each block are randomized. This question-randomization is to avoid that participants learning the order of question types and where to look for answers, which biases participants preference for each mode, instead of reflecting actual term-bank user behaviour or preference. As one participant (no. 2) accurately puts it browsing the stimulus-pair during the experiment: “I just read that!” (*Det har jeg lige læst, det dér!*).

Finally, the blocks (denoted “B1” to “B8”) were randomly ordered throughout the experiment. This block-randomization is to avoid any bias from differences in target term complexity, since target terms are presumably more exposed to the target terms from direct taxation (i.e. B1-B4) are presumably slightly easier compared to indirect taxation (i.e. B5-B8).

4.6 Retrospective interview

After the eye-tracking experiment, I conducted a highly structured retrospective interview (see table 4.3). I used the retrospective interview as one type of auxiliary data to the eye-tracking data. In user testing, retrospection may often involve recording and review of the experiment with the user, while additional questions are asked. It was for technical reasons not possible to replay the experiment to the user, and therefore the retrospection took the form of an after-the-event interview. As part of the debriefing, I conducted a retrospective interview structured around 15 questions, where I asked participants to evaluate the user situation, preference, performance, the perceived difficulty as well as exposure to specialized texts and professional user needs pertaining to concept clarification and term banks.

Table 4.3: Retrospective interview. The topics and questions comprising the retrospective interview. Participants were rating their answers on seven-point Likert-scales (see table E.1 in appendix E for the exact wording of the questions).

Topic	**Question**
User situation	1: How well do you think the experiment resembles a natural user situation (i.e. could you imagine a real situation, where you would have similiar questions)? 2: And how well do you think you got your information need covered (i.e. did the articles and/or diagrams provide you with an answer to the posed questions?
Preference	3: How effective do you think it is to seek information (i.e. find a definition) schematically in an article? 4: How effective do you think it is to seek information (i.e. find a definition) graphically in a diagram?
Performance	5: How well do you think that you performed when you retrieved answers in articles? 6: How well do you think that you performed when you retrieved answers in diagrams?
Difficulty	7: How difficult do you think the specialized text from Statistics Denmark was? 8: Which technical term(s) from the text, do you find the most and least difficult? 9: Which existing technical term(s) from the list, do you find the most and least difficult?
Exposure	10: How often do you read or write Danish specialized texts covering the taxation field in your current job? 11: How often do you read or write English specialized texts covering the taxation field in your current job?
User needs	Imagine a publicly accessible term bank, i.e. a database containing structured information about technical terms from many different subject fields (not only the taxation field). 12: How relevant would a term bank allowing you to search specialized knowledge by means of articles and/or diagrams be for your daily work? 13: Would you prefer to start your search where information is presented in an article, as a diagram, or do you not have special preferences? 14: How often do you need to conduct concept clarification in Danish in your current job? 15:How often do you need to conduct concept clarification in English in your current job?

The retrospective interviews were highly structured with a pre-defined set of questions and responses, where participants were asked to use seven-point Likert-scales to rate (and motivate) their opinion.

"In software evaluation, we can often objectively measure efficiency and effectiveness with performance metrics such as time taken or errors made. Likert scales and other attitudinal scales help get at the emotional and preferential responses people have to the design" (Usability First, 2015). "A structured approach can provide more reliable, quantifiable data than an open-ended interview," (Usability First, 2015). However, an "open-ended" approach allows for an exploratory approach to uncover unexpected information" (Usability First, 2015). Focus-group discussions have the same potential. A focus group is a group of stakeholders, who are brought together to gather input to the design process (Usability First, 2015).

The answers to the retrospective question 1 shows that participants found the experiment valid (mean 4.1 and spread 2.0). Moreover, answers to questions 2, 3-4, 5-6, 10 and 13 are used in the regression approach as explanatory variables of target-user performance (see table 7.1). I disregarded question 11 as it appeared irrelevant for this experiment (mean 2.3 and spread 2.0). The answers to questions 7-9 are used in the reading task (section 6.4). Finally, the answers to questions 12 (mean 4.9 and spread 2.0) reveal a demand among professionals for a term bank. In comparison, regarding answers to questions 14 and 15, the participants reveal a need for concept clarification in their work in Danish (mean 4.0 and spread 1.9), but also in English (mean 3.4 and spread 2.1).

4.7 Focus-group discussion

A few months after the eye-tracking experiments, I held a workshop, where target users drew the concept maps (see section 5.3). Moreover, they were divided into focus groups (named web editors, tax experts, laymen and linguists) and asked to state and evaluate the top-three critical incentives, barriers and effects of having access to a term bank (see table G.2 in appendix G). The most critical answers are outlined in table 4.4.

Overall, the time-saving aspect of gaining access to accurate, correct, reliable and trustworthy information is a clear incentive for the participating target users. Moreover, the effects of the term bank would be to promote consensus and consistency in general, and to provide overview of the tax legislation, in particular, which is necessary when learning new fields or training new staff members.

Table 4.4: Incentives to and barriers against term-bank use. Results of the focus-group discussions. Critical key words on incentives, barriers and effects of term-bank use. In general, each focus group emphasized the potential time-saving effect of having access to a term bank.

Focus group	**Incentives**	**Effects**	**Barriers**
Web editors	Accurate	Tax legislation overview	Know of it
Tax experts	Correct	Train staff	Trust in coverage and up-to-date
Laymen	Reliable and trustworthy	Consensus	Poor-quality answers
Linguists	New fields	Consistency	Complex concept systems

The focus groups were also asked to state critical barriers of term-bank use. It is crucial that the term bank is well-known and that the content is of high quality, without too high complexity, and can be trusted. This dissertation research only pertains to an early phase of the development of a national term bank (DTB-project, 2015). Future work aiming at the establishment of a national term bank in Denmark should integrate the concerns of potential target users in the development as well as meeting the demands for attracting users and promoting the term bank once a first version is ready for implementation.

4.8 Conclusion

Overall, I conclude that mixed methods are necessary for capturing the highly introspective cognitive processes which are activated in target users during the acquisition of terminology and knowledge in an experimental setting (and which will be partly captured by the expertise metrics to be developed in chapters 5 and 6), therefore key to target-user performance.

In addition, I conclude that the most important target users comprise professionals, especially because they exhibit the full scale of expertise throughout the multiple dimensions.

I conclude that the direct measures reflected by self-rating tasks and conventional background variables (participation, motivation, education and exposure) are not fully reflecting expertise. Therefore, the indirect representative tasks (recalling, categorizing, reading and structuring) are developed and evaluated by expertise performance indicators: correctness, speed and depth.

I conclude that the term bank task aimed at target- users' performance in answering multiple-choice questions by consulting the dual-entry modes is well suited for the mixed methods regression models, which provides results on expertise and (local) learning effects.

Finally, I conclude that the retrospective interview and focus-group discussion are qualitative methods to enrich the applied methods.

Chapter 5: Direct expertise measures

In this chapter, I propose direct measures of domain-specific expertise reflected by background variables and self-rating tasks, which contribute to answering the third research theme on expertise.

I provide a brief review of expertise research emphasizing those aspects relevant for the direct measure of the complex domain-specific expertise (section 5.1). The selection of the most important background variables concerning expertise is motivated (section 5.2). Then the self-rating method is motivated as an alternative direct measure of expertise, and the results of the self-rated domain-specific expertise is presented (section 5.3) and compared to the results of the self-rated reference skills (section 5.4).

I expect a set of (explanatory) expertise variables will account for the variance in the proposed direct and indirect expertise measures and in the target-user performance of the dual-entry mode experiments. Therefore, expertise variables are also relevant for the selection criteria, which guide the sampling of the participants.

As will become clear, I chose the regression approach to analyse the self-rating task with the self-rating measure as dependent variable, which allows us to test for significant effects of the expertise variables (section 5.2). However, I chose not to apply the regression approach on the self-rating of reference or English skills, as these measures are show little variance and are therefore analysed descriptively.

5.1 Introduction

A vast body of literature discusses expertise self-rating and characteristics. Fisher (1998) addresses the contribution of training (problem-solving), experience (from rehearsal and practice) and education (acquisition of knowledge through study) on outcomes. Lee, Steyvers, De Young & Miller (2011) show that self-rated expertise measures are outperformed by ranking tasks. Estimates to reflect knowledge by using ranking tasks of e.g. American holidays, avoids the need for relying on self-reported measures. Lee, et al. (2011) use self-report on five-point scales going from no to expert knowledge, and to express their level of confidence.

Bilalić, McLeod & Gobet (2008) suggest paradoxically that experts may fail to solve problems, which require novel approaches, because of their inability to adapt to new demands (expert inflexibility). Moreover, Bilalić, McLeod & Gobet (2008) show that the chosen theories of expertise the retrieval of the familiar solution is supposed to be quick, effortless, and difficult to avoid, therefore (ordinary) experts are prone to use a familiar solution over an optimal solution. However, motivational factors may get the (super) expert to choose the optimal solution (i.e. memory retrieval is followed by a quick search for a potentially better solution).

Experience should be distinguished from deliberate practice (working area), which is the most important contributor to expertise according to Fadde (2009). Cuevas, Fiore & Oser (2002) discuss the discrepancy between self-assessment of performance and actual performance (bias scores) as a measure to ascertain learners' degree of confidence. Expert knowledge is reflected by increased connectedness among critical concepts via graphical representation of conceptual relations (e.g., Glaser, 1989, cited by Cuevas, Fiore & Oser (2002)).

Many studies of expertise focus on the novice-expert difference in knowledge acquisition (see e.g. Alexander, 1992 and Göpferich, 2010) operationalized as (inexperienced) students compared to professionals. I abandon this approach to expertise because of three issues: Firstly, in terminology and specialized language studies, professionals are perhaps the most important target users to include due to their vital role in furthering specialized discourse and promoting the term bank. Secondly, expertise is not a naturally binary variable (expert versus non-expert). On the contrary, expertise is in the abstract notion discussed as at the very least a three-level variable (expert, semi-expert, layperson) (Gouws, 2012). But if participants are left only with these three (artificial) categories in assessing their expertise, I risk missing valuable information (and statistical power) compared to a situation, where I operationalize the self-rated expertise into a numerical, continuous variable. Thirdly, domain expertise or reference skills are usually dealt with in an abstract notion, excluding the metrics assessing the levels, or the variable is dichotomized into a between-group design (professionals versus students) avoiding the problems of assessing any levels. I attempt to measure both domain-specific expertise and reference skills as continuous self-rated variables and include them in the same regression models for performance (see table 7.1).

I define domain-specific expertise as participants' knowledge about the domain stored in their long-term memory i.e. the declarative knowledge (stating terminology) about the particular domain. In addition, domain knowledge also contains procedural knowledge (explaining conceptualizations) (e.g. Alexander, 1992; Kuhn, 2000; Friege & Lind, 2006) i.e. necessary actions for problem solving. I do not include domain-specific (short-term) conditional knowledge (problem-solving), since the application of appropriate problem-solving methods from the taxation domain is not facilitated by term banks, instead information search is likely to influence performance.

As outlined in figure 3.1, I operationalize domain-specific expertise into a continuous variable. To do so, I ask participants to directly rate their knowledge about the taxation domain on a scale going from “very little” to “very high” (see figure 4.2). Consequently, this exercise combines the distinction between declarative, procedural and conditional knowledge into one self-rated measure. In the next chapter (see chapter 6), I ask participants to recall, categorize, read and structure taxation terminology, which I interpret as an indirect measure of the declarative and procedural knowledge. As mentioned, I disregard the domain-specific conditional knowledge, and participants are not asked to solve taxation problems. I do realize that this expertise measure is conditional knowledge dimensions, which are not necessary for high performance. However, I expect the knowledge types to converge, i.e. high conditional knowledge requires high declarative and procedural knowledge. In terms of figure 3.1, domain-specific expertise is evaluated along all three axes.

In addition, I ask participants to directly self-rate their reference skills on the same scale going from “very little” to “very high”, as I expect that this type of non-specific expertise will be highly relevant for target-user performance in the eye-tracking experiments. It should be noted that domain expertise is different from reference skills, which can be defined as participants' knowledge about how to search computer-based references and resources, i.e. more general strategies (Graves, 1996) or problem-solving skills that are independent of the specific domain expertise (Alexander, 1992 and Friege & Lind, 2006). Self-rated reference skills are operationalized in the same way as self-rated domain-specific expertise, i.e. as a numerical, continuous variable.

5.2 Sampling across expertise variables

I chose to follow Kwak (1999) to propose a set of four expertise variables. Kwak (1999) researches factors of knowledge gaps between socio-economic groups and proposes four expertise variables presumably capturing fundamentals of the domain-specific knowledge: Participation reflected by behavioural involvement (section 5.2.1), motivation reflected by issue interest (section 5.2.2), exposure reflected by media use (section 5.2.3) and education reflected by the highest completed level (section 5.2.4). See table 6.4 for participant distribution across the four expertise variables.

5.2.1 Participation

The participants were sampled by means of participation in the domain. In terms of the domain culture framework, intra-domain expertise is reflected by occupation. Therefore, I interpreted the participation as a matter of work place: Half of participants were working at SKAT, while the rest were working outside SKAT, but possibly at work places relevant to the taxation domain. Organizational staff membership does not necessarily entail high domain expertise, and there will also be e.g. taxation experts among participants sampled outside SKAT.

I used the work place as the primary selection criteria. Therefore, 20 participants are staff members of SKAT, while 20 participants are working outside SKAT (see table 6.4).

5.2.2 Motivation

In terms of the domain culture framework, extra-domain expertise is reflected in domain players, i.e. working area. The second expertise variable, motivation, is interpreted as working areas, which presumably reflects expertise better than work place (participation). There will be staff members of SKAT whose work will be entirely unrelated to the issues of the taxation domain e.g. human resource personnel, statisticians, web editors etc. Conversely, there will be staff members outside SKAT, whose work areas will be closely related to taxation. Therefore, I include motivation reflecting issue interest by asking participants to categorize their primary working areas: 17 categories were used (see appendix A) and multiple answers were allowed. Moreover, participants were asked to list typical working tasks in their own words.

5.2.3 Exposure

The third expertise variable is media use, which I transform into exposure to relevant specialized discourse, which is gathered in the retrospection (see appendix E). I ask participants to assess on seven-point Likert-scales, where 1 equals “never” and 7 equals “often” how often they read or write Danish specialized texts from the taxation domain in their daily work. The English counterpart appeared irrelevant because of the low ratings (little variance).

In the retrospective interview, I ask participants to state on a seven-point Likert-scale, how often they read or write Danish specialized texts covering the taxation field in their current job (see table E.1 in the appendix), which I interpret as exposure to the specialized discourse. Answers fall in three large groups: 15 participants never (level 1) read or write specialized texts, 10 participants often read or write specialized texts (level 7), while the remaining 15 lie in between.

5.2.4 Education

The fourth and final expertise variable was education, i.e. the length of their highest completed educational level going from basic education to university degree (i.e. 6 levels, see appendix A). The title of their education was also included. Domain expertise may not necessarily be a question of education, since the career path of similar graduates can turn out very diverse, which I hope to capture by applying other expertise variables.

In the background questionnaire, I ask participants to state the highest level of completed education (see table A.2 in the appendix). Answers fell in two large groups: 29 participants have completed long university educations (i.e. 5 years or more), where as 11 participants have completed tertiary education taking less than 5 years (see table 6.4). In this sample, the 6 levels were not necessary and I dichotomize the variable into two conditions (long and not long). It is the combination of target sample (professionals) and complex domain (taxation), which typically requires long university degrees, and that imposes a bias in the sample.

Age and gender

In addition to the four proposed expertise variables, other conventional background variables are included e.g. age and gender. The latter appear to be a significant predictor of the self-rated domain expertise (see table 5.2), while the former significant in the self-rating of reference skills (see the right panel of figure 5.2).

An unbiased sample of participants should represent the potential target users across the proposed set of expertise variables, as well as age and gender. The latter is easy to sample by, while not all expertise variables exhibit a known value prior to the experiments. 40 volunteers, 23 females (mean age 41.7) and 17 males (mean age 44.0) were sampled from the relevant population of professional potential target users of a term bank.

5.3 Results: Domain-specific expertise

I apply the regression approach for the domain-specific expertise (see section 5.3.2 and see chapter 7.2.3 for a full description of the approach), which I validated (see section 5.3.3), but first we look at the raw data by means of boxplots and scatterplots (see section 5.3.1).

5.3.1 Raw data for self-rated expertise

I begin the statistical analysis by looking at the raw data of participants' answers, which are summarized in table 5.1. Participants rate their domain expertise between 3% to 98% of the total line length with 38.4% mean and 23.2%-point standard deviation. Compared to the two other self-rating exercises comprising reference skills and English language skills, this is a relatively low mean and high spread, suggesting either that taxation experts exhibit low language skills, or an under-estimation bias for rating domain-specific skills.

The mean is the average value, while the spread indicates the dispersion among the observations. A high spread reflects low agreement among participants on a given question, which should worry if we have sampled a homogenous group of participants. High spread is not necessarily a problem, but we should be careful, when we interpret the mean, if it contains many extreme observations. Moreover, high spread will increase the risk of drawing false conclusions in the regression analysis (e.g. failing to reject false null hypotheses).

Table 5.1: Self-rated expertise, reference skills and English language skills. Mean, standard deviation and range for self-rated expertise, reference skills and language (English) skills. Expertise is calculated as the relative line length in percentage of the total line length (80 mm).

All participants (n=40)	Mean	Std. dev.	Range
Self-rated expertise	0.3843	0.2315	0.03 to 0.98
Self-rated reference skills	0.7115	0.1763	0.18 to 0.99
Self-rated English language skills	0.6855	0.1883	0.25 to 1.00

I include age as a potential predictor of the self-rated expertise. Age is potentially a contributor to explaining experience and is therefore included in the analysis below as a numerical variable. It turns out to be non-significant (p-value 0.3221). I propose the four additional expertise variables as explanatory variables (see section 5.2) for self-rating of expertise, i.e. participation, motivation, education and exposure:

Firstly, the majority of participants (72.5%) have completed long-term university education. As mentioned, high spread in the observed data challenges the statistical analysis, because we risk drawing false conclusions. However, the opposite situation of low spread is equally problematic, but for different reasons. The length of completed education shows very little variance, which is not surprising, because the combination of target users (professionals) and domain (complex) is likely to require a long university degree. Unfortunately, explanatory variables with low spread contribute poorly in accounting for the variance in the dependent variable. In other words, the expertise variable education is not useful in explaining the differences in the self-rated expertise.

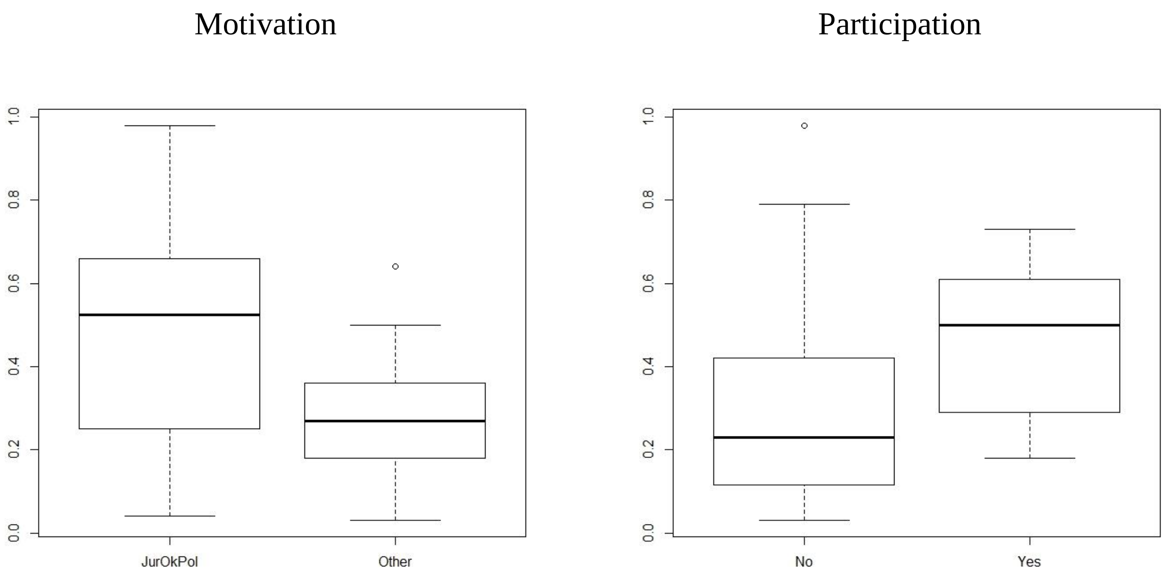


Figure 5.1: Self-rated expertise related to motivation and participation. Boxplot to the left: Self-rated expertise (Expertise_Pct) as a function of the relevant legal, financial or policy areas (JurOkPol) and irrelevant (Other) working area variable. Boxplot to the right: Self-rated expertise (Expertise_Pct) as a function of the work place of participants at SKAT (Yes) or outside SKAT (No).

Secondly, I consider motivation by means of relevant working area. The initial 17 categories were dichotomized into a motivation (working-area) variable with two conditions one assumed relevant to the taxation domain (law, finance, political science) and the other assumed less relevant to the taxation domain (other areas). As shown by the boxplot (see left panel of figure 5.1), relevant working areas produce a higher average self-rated expertise, however, a high spread is following along.

Thirdly, participation interpreted as work place with a relevant taxation authority (SKAT) is included. We use the boxplot (see right panel of figure 5.1) to illustrate graphically the distribution of the observed (raw) data. In the boxplots the median is displayed by a solid line, which is surrounded by a box, where the top/bottom indicates the upper/lower quartiles, and that allows for a comparison between the categories on the horizontal axis. As indicated, participants working at SKAT (Yes) rate themselves higher than participants working outside SKAT (No).

Fourthly, I consider exposure to relevant discourse by asking participants to assess on seven-point Likert-scales, where 1 equals “never” and 7 equals “often,” how often they read or write Danish specialized texts from the taxation domain in their daily work:

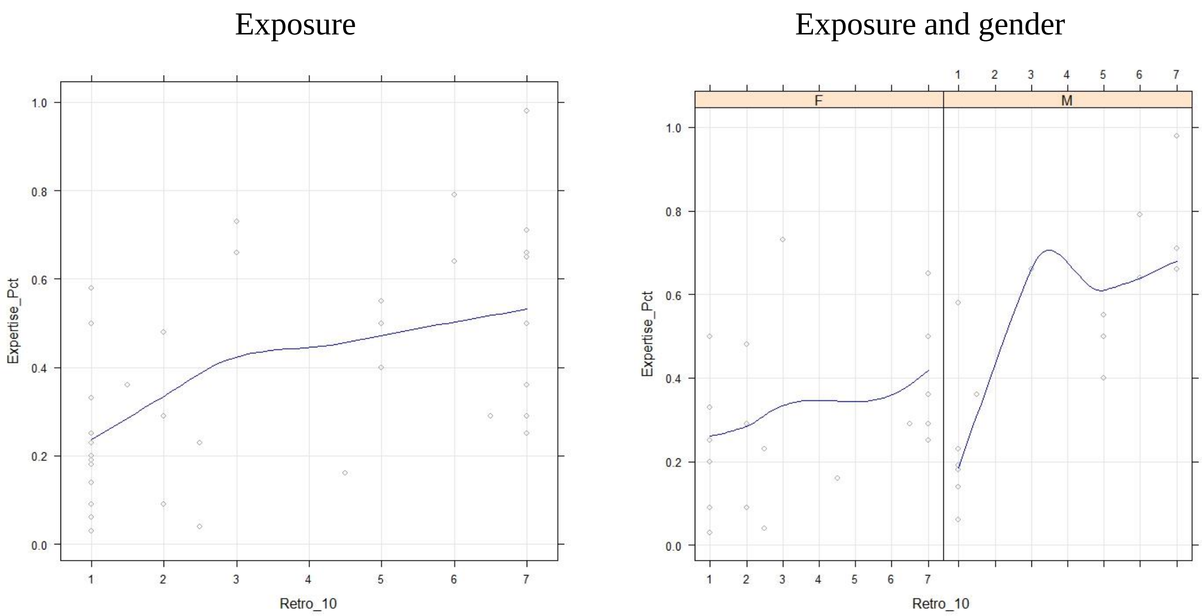


Figure 5.2: Self-rated expertise related to exposure and gender. Scatterplot to the left: Self-rated expertise (Expertise_Pct) as a function of the exposure to Danish of specialized texts (Exposure), which is assessed on a seven-point Likert-scale. Scatterplot to the right: Self-rated expertise (Expertise_Pct) as a function of the use of specialized texts (Exposure: Retro_10) for females (F) and males (M).

Scatterplots are used to display the distribution of (raw) data in the case were data are numerical. The scatterplots in figure 5.2 indicate the effect of the explanatory variable (on the horizontal axis) on the dependent variable (on the vertical axis). Overall, there seems to be a weak positive effect of the use of Danish specialized texts on participants' self-rated expertise (see the left panel of figure 5.2). However, this weak main effect may be concealing significant differences between the sexes (see the right panel of figure 5.2) indicating a statistical interaction between the exposure to specialized texts and gender (see section 5.3.2). It seems that for given levels of exposure to specialized texts, female participants rate their expertise lower than male participants. The regression analysis will reveal whether the interaction is significant.

5.3.2 Regression model for self-rated expertise

A stepwise forward variable selection procedure was used, where we begin with the least interesting variables and end with the most important and in the process only the significant variables are kept in the model. This variable selection avoids the risk of drawing quick, but wrong conclusions.

I build a regression model by including all potentially relevant predictors, in particular, the four different expertise variables and then reducing in a step-wise fashion, reaching simpler and clearer models in which all predictors are significant. A summary of the regression model for self-rated expertise is presented in table 5.2 with the estimated coefficients for the different variables summarised in the second column (denoted "Estimate"), and the associated probability value in the third column (denoted "p(t)").

Table 5.2: Summary of regression model for self-rated expertise. Reference level is female. Exposure reflects the use of Danish specialized texts.

All participants (n=40)		
	Estimate	p(t)
Intercept	0.2363	0.0004
GenderM	-0.0734	0.4392
Exposure	0.0253	0.0735
Interaction: Exposure:Male	0.0603	0.0085

The regression model for self-rated domain expertise shows two significant predictors including one significant interaction between gender and use of specialized texts (exposure). The interaction indicates that the effect of the use of

Danish specialized texts is stronger for males than females (see the right panel of figure 5.2), as the effect is only significant for males (p-value 0.0085), but non-significant for females (p-value 0.0735).

The regression model opens for several interpretations: Firstly, it could be true that females using Danish specialized texts much do not possess a high expertise. Secondly, it could be the case that females, who use Danish specialized texts much, under-estimate their own domain expertise, and/or the corresponding males over-estimate themselves, driving the significant effect for males. Finally, it is, in principle, possible that females are better at (exactly) rating their expertise, while over-estimating their use of specialized texts. It should be noted that only 40 participants have rated their expertise, which generates a limited number of observations and hence a potential for improving the statistical power of the analysis. High statistical power means that false null hypotheses are correctly rejected avoiding false negatives, which allows for the detection of significant effects explaining the variance in performance.

5.3.3 Validity of self-rated expertise

An effect of the experimental context is the self-rating nature of the task, which in itself and independent of differences in the actual knowledge levels, will show variance not reflecting difference in the knowledge levels, but express an element of confidence leading more confident participants to slightly over-estimate own skills and less confident participants to slightly under-estimate themselves. In particular, the chosen domain is taxation, which comprises several domain cultures (see chapter 3), possibly eroding participants' confidence leading to an under-estimation bias.

Another context effect adding to under-estimation bias is that the question turned out hard to grasp for the participants. Already during the pilots, participants voiced this problem by wondering about the rating of the narrow case where you know a lot about a small sub-domain versus the broad case where you have knowledge throughout the whole domain. In the briefing of participants upon arrival, I did specifically ask them to follow their immediate intuition, instead of trying to infer or calculate their average level of knowledge.

A simple way of overcoming some of the under-estimation bias could be to rephrase the question from the perhaps too general “How large is your knowledge in the taxation domain” to something more specific e.g. “How large is your knowledge in the part of the taxation domain that is the most relevant to you.” If you are an expert of a sub-domain, you will presumably be a semi-expert of other parts of the domain, but that should not erode your confidence or the expert knowledge of your sub-domain.

It will add to the validity of the self-rated expertise how the variable accounts for the variance in performance in the subsequent experiments, i.e. the expert performance analysis (see chapter 6) and the dual-entry-mode experiment (see chapter 7). However, participants seem able to perform well without applying much domain expertise, which is realized by the non-significance of self-rated expertise in the following analyses.

5.4 Results: Reference skills

I expect that the skill most crucial to participants' performance in the dual-entry mode experiment is the ability to apply general strategies of information search. In the experiment, the article questions are indeed a matter of reading a table of terminology information and comparing that with the available answers, which very much resembles searching term banks or other reference works, and therefore I also ask participants to self-rate their reference skills, i.e. the ability to search information electronically. In the diagram questions participants must also be able to infer knowledge from the diagrams and compare that to available answers, which is resembling information search to the same extent.

Participants are asked to assess their reference skills in the same manner as domain knowledge on a line (80 mm) starting at “very little” and ending at “very high”, and with an indicator at the middle (40 mm). The exact wording of the question was: “How good are you at seeking information electronically? Please put a mark on the line below” (see table A.2 in appendix A). I expect that self-rated expertise (specific to the domain) and reference skills (generic or non-specific to the domain) are independent of each other, because there is no reason to believe that reference skills will contribute to explain any of the variance in the self-rated domain expertise, which is confirmed in the left panel of figure 5.3 which shows no clear connection between the two variables.

Domain expertise Age

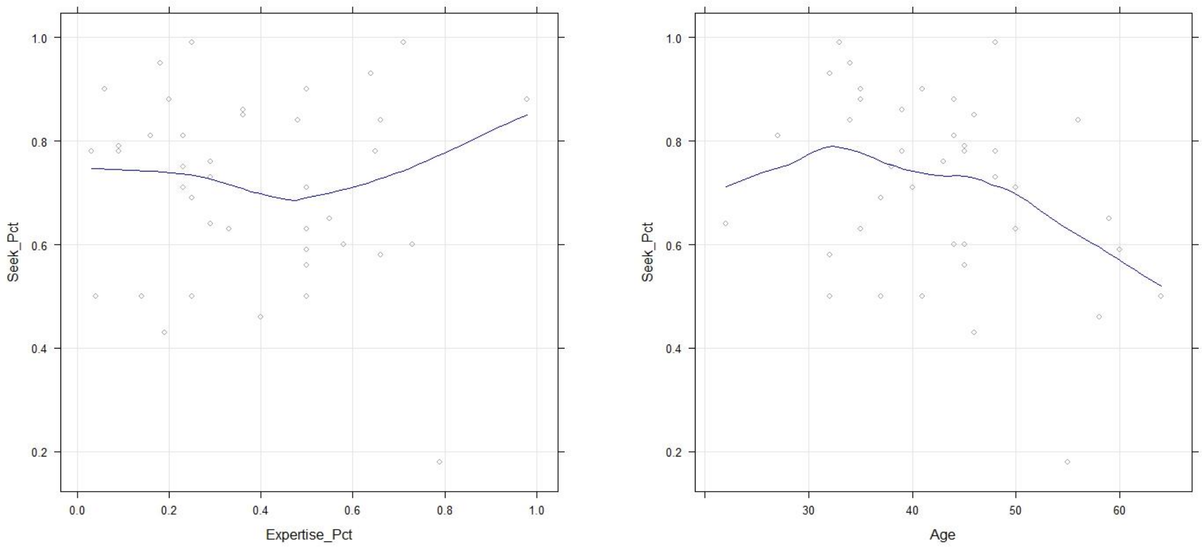


Figure 5.3: Self-rated reference skills related to self-rated expertise and age. Scatterplot to the left: Self-rated reference skills (Seek_Pct) as a function of self-rated domain expertise (Expertise_Pct). Scatterplot to the right: Self-rated reference skills (Seek_Pct) as a function of age.

Moreover, age shows a weak negative effect in explaining the variance of self-rated reference skills (see the downward slope of the right panel of figure 5.3), suggesting that older participants rate their search skills lower compared to younger. I encountered no similar age effect on the self-rated domain expertise.

5.4.1 Self-rated reference skills

The self-rated reference skills range from 18% to 99% around 71.2% mean, which is relatively high compared to self-rated domain expertise (38.4%) and with a lower spread 17,6%-point (compared to 23.2%-point) (see table 5.1). Hence, there seems to be no corresponding under-estimation bias in the self-rating of reference skills.

Moreover, the higher mean and lower spread of self-rated reference skills (see the right panel of figure 5.4) compared to self-rated domain expertise (see the left panel of figure 5.4) make the gender differences in the reference skills less clear (see right panel of figure 5.4).

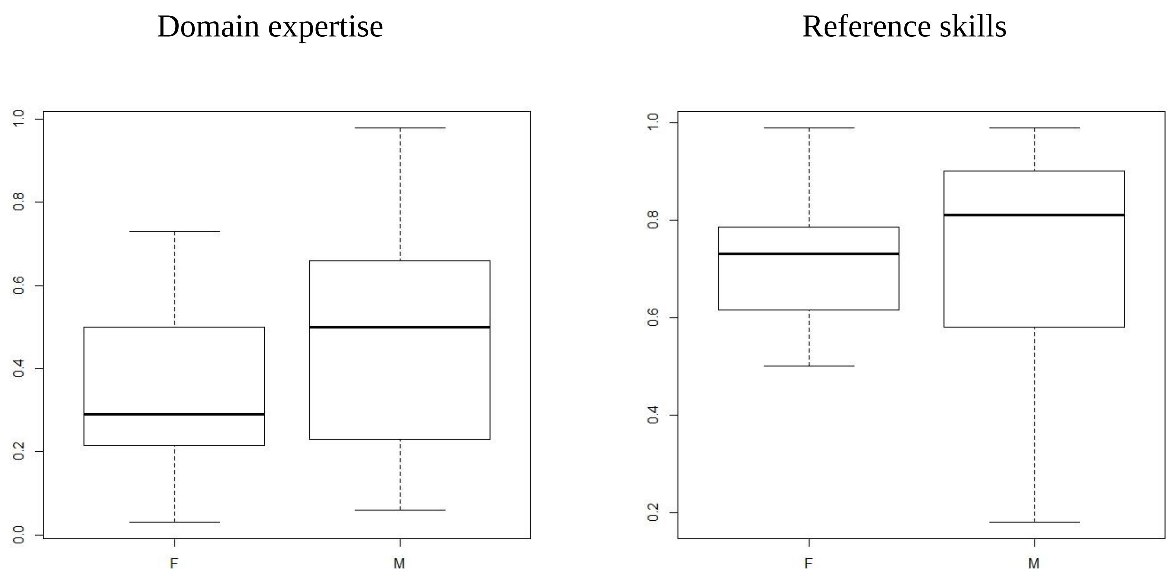


Figure 5.4: Self-rated expertise and reference skills related to gender. Boxplot to the left: Self-rated expertise (%) as a boxplot function of gender (F=Female and M=Male). Boxplot to the right: Self-rated reference skills (%) as a boxplot function of gender.

The self-rated reference skills and the self-rated English language skills are not subject to regression analysis because they are less relevant for explaining the performance in the dual-entry mode. In addition, they exhibit less variance.

5.4.2 Used and non-used reference tools

In addition to the self-rating task, participants were asked about their use of seven listed reference works (see table 5.3). The use of reference works was captured by a survey-like question quantifying the use and/or knowledge of existing reference works. The question sounded: “State how many times on a typical working week, you use electronic search tools.” The tools listed were chosen to capture encyclopedic (top-three items), terminology and dictionary (bottom-three items) tools (see table A2, section 10, of Appendix A). Unfortunately, many zeros appeared (see table 5.3) and Google is the only search tool, which is rated above zero by all participants and therefore capturing a potential reference skill.

Table 5.3: The use and knowledge of reference works. “0a” denotes “0 times, since I do not *know* the tool.” “0b” denotes “0 times, since I do not have *access* to the tool.” “0c” denotes “0 times, since I do not *need* the tool.” “0d” = “0 times, since I do not *like* the tool.” “1” covers 1-10 times. “2” covers 10-50 times. “3” covers 50-100 times. “4 covers More than 100 times. *One participant missed the last question.

All participants (n=40)

Tools:	0a	0b	0c	0d	1	2	3	4	Total
Google					7	17	6	10	40
Wikipedia			12	1	18	6	3		40
Denstoredanske.dk	10	3	18	1	8				40
Iate.europa.eu	25	1	9		5				40
Ordbog.gyldendal.dk	6	2	21	1	3	5	1	1	40
Ordbogen.com	6	4	17		11	1	1		40
Oxforddictionaries.com	12	1	18	1	3	4			39*

The possibility of having many zeros was anticipated and alleviated by grading the zeros into categories of reasons for not using the particular tool. Thus, participants were able to qualify a zero by adding a reason – that they don"t “know”; have “access” to; “need”; or “like” a particular tool.

Generally speaking, only very few participants disliked the listed tools, and rare use of a tool is generally a matter of not needing or knowing a tool. Among the encyclopedic tools, 32 participants (80.0%) state, they do not use “Denstoredanske.dk” on a weekly basis, a national reference work aiming at being Danish internet users' preferred source of information (Gyldendal, 2015a). From this, 18 participants (56.3%) state they do not need the tool. For the terminology tool, i.e. the multilingual term base of the EU (IATE, 2015), 35 participants (87.5%) state they do not use the tool “Iate.europa.eu” on a weekly basis, and from this 25 participants (71.4%) say they do not know the tool. Among the dictionary tools, the two bi-lingual, “Ordbog.gyldendal.dk” (Gyldendal, 2015b) and “Ordbogen.com” (Ordbogen, 2015) are not used on a weekly basis by 30

(75.0%) and 27 (67.5%) of participants, and the main reason is that they do not need the tool. For the last (multi-lingual) English online dictionary, participants also state that they do not know the tool “Oxforddictionaries.com,” which is an online dictionary of British and world English (Oxford Dictionaries, 2015).

Apart from stating their weekly use of the listed search tools, participants were able to state non-listed (both electronic and printed) tools. The answers included other mono- or multilingual dictionaries (e.g. German language), domain-specific knowledge bases (e.g. legal, medical), business software (e.g. statistical, financial, intranet), news sites and parliament sites.

5.4.3 Validity of self-rated reference skills

It seems easier for the participants to estimate their reference skills than their domain-specific expertise reflected by higher mean and lower standard deviation. Perhaps, this is due to the nature of the question, which is conceptually easier to grasp, since we all seem to know if we are good at searching information. In addition, no marked differences across the sexes appear in self-rating reference skills. What should be remembered is that participants are all professionals. The professionals of the sample are all having working tasks which entail navigating and searching information electronically, and therefore the self-rated reference skills might turn out quite precise.

5.5 Conclusion

I conclude that the set of expertise variables crucial to the performance of target users comprise participation (work place), motivation (work area), education and exposure to the specialized discourse. In addition, I use the participation and partly the work area as selection criteria for sampling the 40 participants for the dual-entry mode experiments.

I conclude that it is reasonable to measure complex domain-specific expertise in one dimension, because I expect the domain-specific knowledge (i.e. declarative, procedural and conditional) to converge. In terms of figure 3.1, the assessment is in principle reflecting all three dimensions. I would have expected reference skills (and perhaps also English language skills) to influence performance, however, a regression analysis was discarded because of very little variance in the dependent variables.

I conclude that the self-rated domain expertise (i.e. self-rating knowledge and practice) in the taxation domain produced an under-estimation bias compared to the self-rating of reference skills. Only one of the proposed expertise variables, exposure to relevant specialized discourse, produced a significant main effect and interaction with gender on self-rated expertise. Indeed, I encounter a risk that the simple self-rated measure applied on the complex taxation domain might be invalid.

Regarding the research theme, I conclude that proposed direct measures of expertise such as self-rating challenge participants, who show confusion and tend to under-estimate their expertise. Therefore, it seems necessary to propose indirect measures, where participants are not directly revealing (and possibly underestimating) their expertise.

Chapter 6: Indirect expertise measures

In this chapter, I propose indirect measures of domain-specific expertise by means of representative tasks, i.e. recalling, categorizing, reading and structuring, with associated expert performance indicators (correctness, speed and depth). Chapter 6 contributes to answering the third research theme on expertise.

I review the literature on expertise to motivate the choice of relevant representative tasks as well as key indicators of expertise performance (section 6.1). The evaluation of each task by means of three performance indicators is discussed separately for the recalling task (section 6.2), the categorization task (section 6.3), the reading task (section 6.4) and the structuring task (section 6.5).

I chose the regression approach to analyse the categorization task with correctness as dependent variable. However, no clear dependent variable can be defined for the remaining representative tasks (recalling, reading and card-sorting), and that is the reason why I chose to investigate the expert performance indicators descriptively.

6.1 Introduction

Eye tracking has also been used for showing differences between expert and novice user (Bednarik, Kinnunen, Mihaila & Fränti, 2005). Lachner, Gurlitt & Nückles (2012) show graph-oriented measures by detecting differences in expert and intermediate structures, where knowledge encapsulation, i.e. the more omission of concepts and the shorter inference path, and knowledge integration, i.e. expert explanation less fragmented, but connectedness is not significant.

Mason & Singh (2011) show a connection between categorization and expertise. Moreover, it is shown that initially experts represent the problem at a more abstract (i.e. context independent) level compared to novices, who focus on surface features. Friege & Lind (2006) test for declarative knowledge by asking about definitions, laws, examples, magnitudes, etc. In addition, a concept mapping task is aimed at evaluating the interconnectedness, hierarchisation and the level of abstraction of declarative knowledge. The graphs are evaluated by means of graph-theory, reference maps and relations-evaluation. As discussed by Fadde (2009), often experts do not perform better than non-experts on representative tasks. Either the identified experts are not actually experts, or the identified tasks are not really representative. The question of expertise is relative, but experience is essential.

The review on expertise research does not constitute a general review on the psychological literature on expertise but rather a brief review of the most important tasks. Following Charness & Tuffiash (2008), who introduce representative tasks to demonstrate the superior performance of experts, I propose four representative tasks relevant to domain-specific terminology and knowledge:

Firstly, an open-ended, association task of recalling taxation terms (see section 6.2) focusing on the terms (expressions) of the domain-specific terminology. Secondly, a categorizing task distinguishing existing taxation terms from non-existing pseudo-terms (see section 6.3) focusing on the concepts (meaning) of the terminology presented. Thirdly, a reading task using an authentic, specialized text containing long and complex, domain-specific terms (see section 6.4) focusing on the discourse, in which the terminology is occurring. Fourthly, a card-sorting task focusing on the conceptual structures (see section 6.5). I develop relevant performance indicators, inspired by some of the expert characteristics discussed by Rikers & Paas (2005). I evaluate whether expert performance in each of the representative tasks corresponds to the expert characteristics (correctness, speed and depth).

6.2 Results: Recalling terms

In the recalling task (see section 4.4.1), correctness measures the number of correct terms on participants' list of ten terms. As mentioned speed had to be discarded, because it was necessary to minimize the time in front of the eye tracker and the response time was not recorded when participants completed the background questionnaire prior to the dual-entry mode experiment. In addition, it should be noted that the recalling task becomes merely an indirect indicator of participants' depth, because they were asked about the first ten terms that came to mind, not the most deep ones, and it is possible that participants with deep (expert) knowledge were keener on completing the list quickly, than on exhibiting depth.

Writing the first ten taxation terms that come to mind turned out to be less demanding for the participants than expected (see section 6.3.1). However, the list adds to our understanding of the limitations of the complexity indicators that are applied in reading research of generalized language which cannot be directly applied in specialized language research and that challenges the development of a depth indicator (see section 6.3.2).

6.2.1 Correctness: Term lists

Each of the 40 participants were able to write at least seven taxation terms, while 70% were able to produce a complete list of the required ten taxation terms (see table 6.1).

Table 6.1: Recalling ten terms. Number and share of participants, who were able to write a list of up to ten terms from the taxation domain. The 40 participants listed in total 377 terms, 177 unique terms.

All participants (n=40)		
Term list	Number	%
1. term	40	100.0%
2. term	40	100.0%
3. term	40	100.0%
4. term	40	100.0%
5. term	40	100.0%
6. term	40	100.0%
7. term	40	100.0%
8. term	37	92.5%
9. term	32	80.0%
10. term	28	70.0%

I evaluated all stated terms as correct, because it turned out very difficult to reject terms as not being domain-specific terms from the taxation domain. Obviously, some terms are more fundamental to the domain than others (see discussion in section 3.3), but it is almost impossible to reject that a term on the list is not designating a concept within the domain even though it may seem quite peripheral. At first glance, a simple remedy of the large term lists (few blanks) could have been to increase the number of required terms, which was avoided due to the need for keeping the overall completion time of the background questionnaire acceptable. In future research, it is recommendable to design the task to challenge participants further by increasing the number of terms (and hence observations) and include a time constraint, i.e. asking participants to list as many and as complex taxation terms as possible e.g. within ten minutes.

Almost half of the listed terms, i.e. 177 out of the total 377 recalled terms (47.0%) are unique terms, and 115 terms out the 177 unique terms (65.0%) are only mentioned by one participant. Each term is listed by up to 17 participants, i.e. the term “top-bracket tax” (*topskat*) is listed by 17 participants, the term “land tax” (*ejendomsskat*) is listed by 12 participants, the term “property value tax” (*ejendomsværdiskat*) is listed by 11 participants, etc. Terms mentioned by five participants or more are shown in table 6.2.

Table 6.2: Terms mentioned by five participants or more. “Sample frequency” is the frequency in the sample. “Length” is measured by characters (including spaces). The term length ranges from 4 to 34 characters, with a mean score at 13.0027. “Word frequency” is measured as the term's frequency in KorpusDK (generalized language corpus).

All participants
(n=40)

Term answer	Sample frequency	Length	Word Frequency
topskat	17	7	85
ejendomsskat	12	12	261
ejendomsværdiskat	11	17	145
ligningsmæssige fradrag	10	23	33
moms	10	4	1,169
bundskat	9	8	114
fradrag	9	7	1,039
personfradrag	9	8	148
årsopgørelse	8	12	13
forskudsopgørelse	7	16	33
B-indkomst	6	10	0
befordringsfradrag	6	18	89
kirkeskat	6	9	106
rentefradrag	6	12	565
trækprocent	6	11	117
A-skat	5	6	0
arbejdsmarkedsbidrag	5	20	174
grundskyld	5	10	81
restskat	5	8	120

I became aware of the weakness arising from the fact that the terms appearing in the introduction to concept clarification (see table B.1 in appendix B) might have biased performance, and that participants might have looked below the ten lines ready for the listing into the subsequent categorization task (see table B.2 and B.3 in appendix B), the latter potentially exposing participants to the terms “land tax” (*ejendomsskat*) and “middle-bracket tax” (*mellemskat*) and to the pseudo-term “value-subtracted tax” (*mindreværdiafgift*): The first one occurs second in table 6.2 and might have caused associations to the semantically closely related “property value tax” (*ejendomsværdiskat*) occurring third. The second one might have caused associations to the first in table 6.2 as the term “middle-bracket tax” (*mellemskat*) and “top-bracket tax” (*topskat*) are semantically closely related. Despite the pseudo-status of the third one, it might have lead participants to think of “value-added tax” (*merværdiafgift*) synonymous with “value-added tax” (*moms*) ⟐ occurring fifth in table 6.2.

6.2.2 Depth: Frequency of domain

The high correctness performance (and missing data on speed) necessitates the third performance indicator, depth. Let us evaluate the recalling task by applying the depth indicator, i.e. I apply a depth analysis to the term list from the recalling task to see if experts demonstrated superior performance.

The depth indicator must reflect the problem representation, where experts' representation is deep and less superficial compared to novices (Rikers & Paas, 2005). In domain-specific terminology, depth is different from complexity, which comprises objective criteria connected to the morphology of the term, often including frequency, length and predictability. The lengths (measured in characters) range from 4 to 34 rather randomly across the list (see table 6.2), and the word frequencies (measured in a generalized language corpus) of the listed terms range from 0 to 1,169 equally without any clear pattern. Therefore, I need to abandon those two complexity indicators and develop another depth indicator for domain-specific terminology. Depth is not related to the linguistic features of terms, it is related to the semantics, i.e. the concepts designated by the terms in question. Based on this, I propose that the depth should be qualified by assessing the conceptual contents of the term in question. I operationalize the indicator of depth by assessing in the relevant specialized discourse, which is chosen to be SKAT's legal instructions (SKAT, 2014) constituting a relevant and domain-specific legal corpus, where the frequencies are used to measure depth:

A high frequency indicates that many legal documents concern this term i.e. a broad (superficial, less deep) term, while a low frequency indicates that few legal documents cover this term, i.e. a narrow (deeper) term. The frequencies are then categorized, i.e. the depth is measured by means of frequency bands ranging from 1 to 7 representing (approximately) equally sized portions of the (unique) term-list members in each band (ranging from 20 to 30) producing term-list densities (tokens) ranging from 29 to 68 (see table 6.3).

Table 6.3: Depth performance indicator of recalling task. Frequency bands assessed by means of SKAT's legal instructions and used as indicators of depth ranging from 1 to 7. 177 out of the total 377 recalled terms are unique terms. Mean depth of term lists is 3.7215 with standard deviation 1.8070.

All participants (n=40)

Depth	Frequency bands	Term distribution in each band	Term-list density in each band
1	Above 1,000	21	48
2	500-999	27	65
3	200-499	29	68
4	100-199	30	60
5	50-99	20	64
6	10-49	28	43
7	1-9	22	29
	Total	177 (unique)	377 (total)

The depth indicator was expected to demonstrate superior expert performance, however, we may ask ourselves if experts identified by direct measures exhibit depth performance in the recalling task. Depth performance does not seem to vary much across the different levels of the proposed expertise variables education, participation, motivation (see table 6.4) and exposure (see figure 6.1), although the depth mean is slightly higher for participants with a long education (3.8199), participants working with relevant areas (3.9538), and also for the participants working at SKAT (3.7760). This does not necessary constitute a problem, as the expertise variables were slightly biased (see section 5.2)

Table 6.4: Depth performance indicator and expertise variables. Depth mean and standard deviation across the expertise variables "education" is measured by length, "participation" is measured by work place, and "motivation" is measured by working area. The number of participants and terms (observations) in each category of expertise variable is stated.

All participants (n=40)

	Mean	Std. Dev.	Participants	Number of listed terms
Education				
- 5 years or more	3.8199	1.8527	29	272
- Below 5 years	3.4667	1.6645	11	105
Participation				
- SKAT	3.7760	1.7891	20	192
- Outside SKAT	3.6649	1.8286	20	185
Motivation				
- Relevant	3.9538	1.8980	18	173
- Irrelevant	3.5245	1.7061	22	204
Total	3.7215	1.8070	(40)	(377)

Exposure to the relevant specialized discourse, i.e. the use of Danish specialized texts, seems to have no significant effect on the proposed depth indicator (see the left panel of figure 6.1). Moreover, the self-rated domain expertise does not affect the depth indicator (see the right panel of figure 6.1). In other words, depth as an objective indicator of expert performance does not seem to be affected by the self-rated expertise (the subjective indicator of expertise).

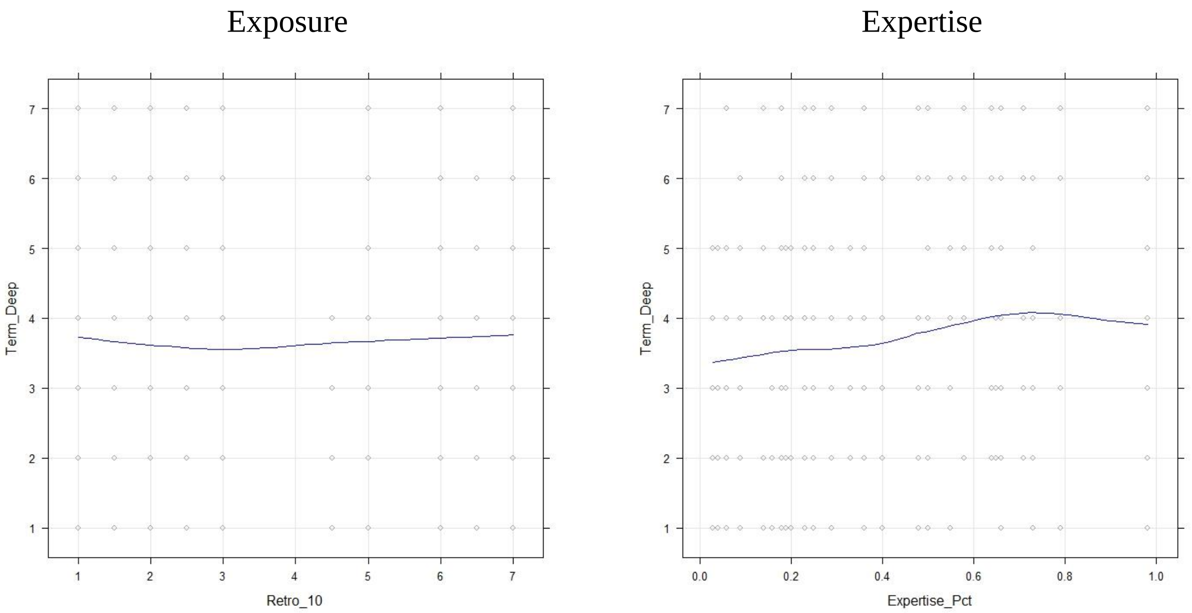


Figure 6.1: Depth of recalling task related to exposure and self-rated expertise. Scatterplot to the left: Depth (Term_Deep) as a function of the exposure variable, use of Danish specialized texts (Exposure: Retro_10). Scatterplot to the right: Depth (Term_Deep) as a function of self-rated expertise (Expertise_Pct).

As mentioned, participants were not directly asked to list the deepest (or most complex) terms from the taxation domain, instead they listed the first terms that came to mind (without any time constraints), which may weaken the depth analysis by distorting the frequency bands producing inadequate depth indicators, which are then reducing the importance of expertise. Ideally, other depth indicators should be included to fully reflect the conceptualizations and problem representations of the taxation domain.

6.3 Results: Categorizing terms and pseudo-terms

For the categorizing task (see section 4.4.2), correctness is analysed by choosing a (binary) regression approach for the dependent variable correctness, while including as explanatory variables the proposed expertise variables (see section 5.2). In addition, a ranking of the definitions produced by participants is used as the depth indicator. As mentioned, response times were not recorded and the speed indicator is therefore unavailable.

6.3.1 Correctness and significant predictors

The main hypothesis for this analysis is that experts make fewer errors (Rikers & Paas, 2005). The terms and pseudo-terms were presented in random order. In particular, participants were presented with a randomized list of ten existing taxation terms, which were contaminated with five semantically plausible pseudo-terms and asked to state, whether they existed or not. If participants believed an item existed, they were asked to give a short description of the meaning (see table B.3 in appendix B). Therefore, correct answers are "yes" for existing terms and "no" for pseudo-terms, see table 6.5.

Table 6.5: Terms and pseudo-terms of the categorization task. Trials are ordered by existence and correctness scores. "Existence" has two conditions: "Yes" for existing, and "No" for non-existing. "Correctness %" is the share of correct answers. In other words, correctness in the upper half, corresponds to stating (correctly) that terms exist, while correctness in the lower half, corresponds to stating (correctly) that pseudo-terms do not exist.

All participants (n=40)

	Existence	Correctness %
grøn afgift	Yes	97.5
energiafgift	Yes	97.5
skatteprovenu	Yes	97.5
ejendomsskat	Yes	92.5
mellemskat	Yes	85.0
virksomhedsskat	Yes	82.5
personlig indkomstskat	Yes	75.0
skatteindtægt	Yes	70.0
forbrugsbegrænsende afgift	Yes	62.5
fiskal skat	Yes	42.5
C-skat	No	97.5
kaffemoms	No	95.0
mindreværdiafgift	No	95.0
vindafgift	No	87.5
forskudsfradrag	No	80.0

The average correctness of existing terms (80.25%) is lower compared to the non-existing pseudo-terms (91.0%), suggesting that participants were better at categorizing the latter. Correctness ranges from 42.5% to 97.5% for the terms and

80.0% to 97.5% for the pseudo-terms. Thus, the pseudo-terms turned out to be rather obvious to participants. In contrast, in the retrospection, participants stated their surprise by the bottom-five existing terms, which seemed unlikely to participants due to the seemingly informal and non-legal status, i.e. “tax receipts” (*skatteindtægt*) seemed too informal compared to the formal “tax revenue” (*skatteprovenu*), despite the fact that they are synonyms, and that resulted in only 70.0% correct answers (i.e. 30.0% incorrectly categorized the term as non-existing).

A summary of the regression model for the dependent (binary) variable correctness is presented in table 6.6.

Table 6.6: Summary of regression model for correctness in the categorization task. Reference level working area is the relevant condition (legal, finance, or policy). “Exposure” is reflected by use of Danish specialized texts, while “motivation” is reflected by working area.

All participants (n=40)	Estimate	p(t)
Intercept	1.3520	0.0059
Exposure	0.2043	0.0136
Motivation (Irrelevant)	0.7052	0.1226
Interaction: Motivation:Exposure	-0.2601	0.0139

Only one of the proposed expertise variables, namely, exposure to relevant discourse, measured by means of the self-stated degree of use of Danish specialized texts has an overall significant effect (p-value 0.0136) on correctness (see table 6.6). The estimated coefficient is positive reflecting that, overall, increased exposure to relevant discourse increases the correctness. However, the exposure variable interacts significantly (p-value 0.0139) with the motivation variable (working area), i.e. the exposure effect is significantly different (and in fact absent) for the irrelevant working area on correctness compared to the relevant, where it has a facilitatory effect. Hence, the correctness analysis demonstrates superior correctness performance of participants exhibiting high exposure and motivation. The remaining two expertise variables (education and participation) as well as the self-rated expertise were non-significant. In other words, the direct measures of expertise are only in two cases relevant for the representative task, categorization.

6.3.2 Depth: Detail of definitions

While running the experiments, I got the impression, that participants did think very carefully about the existence of the 15 presented terms and pseudo-terms and also did their best at trying to define (in their own words) the terms they believed existed. However, the quality of the definitions varied much, which made a reliable ranking of the answers impossible.

Participants produced a high depth variance (differences in levels of detail and precision) on the definitions that they offered from the precise: “the tax which according to the act on corporation taxation is collected as an interim tax on the profits saved, at the moment 25 per cent” (*den skat som iht. Virksomhedsskatteloven opkræves som en acontoskat på det opsparede overskud pt. 25%*) to the vague and unprecise “companies' tax” (*virksomheders skat*). Moreover, I encountered many blanks, which occur due to the experimental design, where participants only define, what they categorize as existing terms, i.e. they are not defining terms, which they (incorrectly) may have categorized as non-existing. In addition, participants have an option of avoiding the definition by answering “Yes, the expression exists, but I do not know the meaning.” It should be noted that it is not fruitful to separate the definition task from the categorization, as it made participants carefully consider their answer. But further developments of the task avoiding blank options and asking for more precise meanings are likely to improve the quality of answers.

6.4 Results: Reading terms in an authentic, specialized text

In the reading task (see section 4.4.3), correctness is obviously not a meaningful performance indicator of primary interest, but I do include participants' own perception of the difficulty of the contents (see table 4.4) as an indicator of participants’ correct understanding. The evaluation of correctness would improve by direct follow-up questions on participants understanding of the content. Since the reading task is primarily aimed at measuring the participants' processing time, it was necessary to conduct this task in front of the eye tracker recording the reading time of the two sections of different levels of depth (AOILeast and AOIMost). It should be noted that the purpose of the reading was not stated directly, rather it was part of building the scenario for the following dual-entry experiment (also executed in front of the eye tracker).

Participants read an authentic, representative, specialized text from the taxation domain, containing taxation terms as well as terms from the public finance domain alongside non-specific words (see figure C.4 in appendix C). Two sections of the text were chosen as comparable areas-of-interest (AOIs) for the processing time analysis, where I expect experts to process the most complex section faster (AOIMost).

6.4.1 Speed: Processing time of depth

Following Jensen (2009), I distinguish between the (objective) text complexity (measured by proposed complexity indicators) and the (subjective) text difficulty (measured by the processing time of readers). In the absence of a large domain-specific corpus, I measure frequencies of terms in a general-language corpus, which will result in relatively low, perhaps even irrelevant, frequencies without much relation to complexity (see table 6.2 for word frequency measures of terms). Moreover, the length of terms in characters seems a poor indicator of complexity, as the most frequently recalled terms are relatively long e.g. “tax relief” (*ligningsmæssige fradrag*) ⟐ or “labour-market contribution” (*arbejdsmarkeds-bidrag*) (see table 6.2). Therefore, I abandon word frequencies and term lengths as indicators of complexity, although these indicators are applied in reading research on generalized language (e.g. Rayner, 2009).

Instead, I focus on the processing time of two equally sized sections of the chosen specialized text (see figure C.4 in appendix C), which contain domain-specific terms from the taxation domain and the public finance domain: Hence, the most deep (denoted “AOIMost”) section begins with “Some types of taxes...” (*Nogle skatter og afgifter…*) and ends with “... consumption-limiting duties” (*forbrugsbegrænsende afgifter*); contains eight taxation terms concerning duties; five public finance terms concerning economic behaviour; and has a LIX-score of

56 (Jensen, 2009). The least deep (denoted “AOILeast”) section begins right after the most complex with “For many years, total taxes and duties have...” (*De samlede skatter…*) ⟐ and ends with ”... income from value-added taxes” (*...personlige indkomstskatter*); contains six taxation terms concerning revenues; five public finance terms concerning the national accounts; and has a LIX-score of 45 (Jensen, 2009). Moreover, this approach allows us to discard the first lines of the text (where participants are getting started), as well as the last line (where participants are getting concerned about continuing the experiment).

The processing time includes not only fixations, but also saccades (i.e. movements between fixations) recorded inside an AOI, since cognitive processing may take place during rapid eye-movements (Holmquist et al., 2011). Processing time differs across the two sections with the longest mean time (22,567 ms) on the least deep section and the shortest mean time (20,522 ms) on the most deep section (see table 6.7).

Table 6.7: Processing of the reading task. Mean and standard deviation of reading time (ms) in the reading of the most and least deep sections. The eye-tracking data of six participants had to be discarded due to technical problems. Processing times of the text (AOI) includes dwell time (i.e. the sum of durations from all fixations and saccades that hit the AOI) and the duration of saccade entering the AOI. Mean per word is the mean processing time divided by number of words in the sections.

Number of participants (n=34)			
	Mean	Std. Dev.	Mean per word
Least deep (AOILeast)	22,566.9	10,801.8	376.1
Most deep (AOIMost)	20,522.3	8,737.2	331.0

It is possible that this overall, and counter-intuitive, difference in mean processing time across the two sections covers highly deviating processing times across participants. Therefore, we would expect expertise effects to occur in the most complex section, in particular, with experts showing superior (fast) performance, while differences across participants would be less strong in the least deep section, where expertise effects are expected to be less important. Considering the most deep section, we do see that the processing time mean is shorter for long-term university educated (20,130 ms) participants working with relevant areas (18,352 ms), while longer processing time occurs for the participants working at SKAT (23,077 ms) (see table 6.8).

Table 6.8: Processing and expertise variables. Reading time (ms) of the *most* deep text (AOIMost) across the expertise variables: "Education" is measured by length, "participation" is measured by work place, and "motivation" is measured by working area. The eye-tracking data of six participants had to be discarded due to technical problems.

All participants (n=34)	Mean AOIMost	Std. Dev. AOIMost
Education		
- 5 years or more	20,130	6,491
- Below 5 years	21,421	12,438
Participation		
- SKAT	23,077	11,172
- Outside SKAT	18,672	5,790
Motivation		
- Relevant	18,352	6,841
- Irrelevant	21,921	9,523
Total	20,522	8,737

In addition, the exposure to relevant specialized discourse, i.e. the use of Danish specialized texts (Exposure), seems to have no significant effect on the processing time of the most deep section (see the left panel of figure 6.2), which also seems to be the case for the self-rated domain expertise (Expertise_Pct) (see the right panel of figure 6.2).

Exposure Expertise

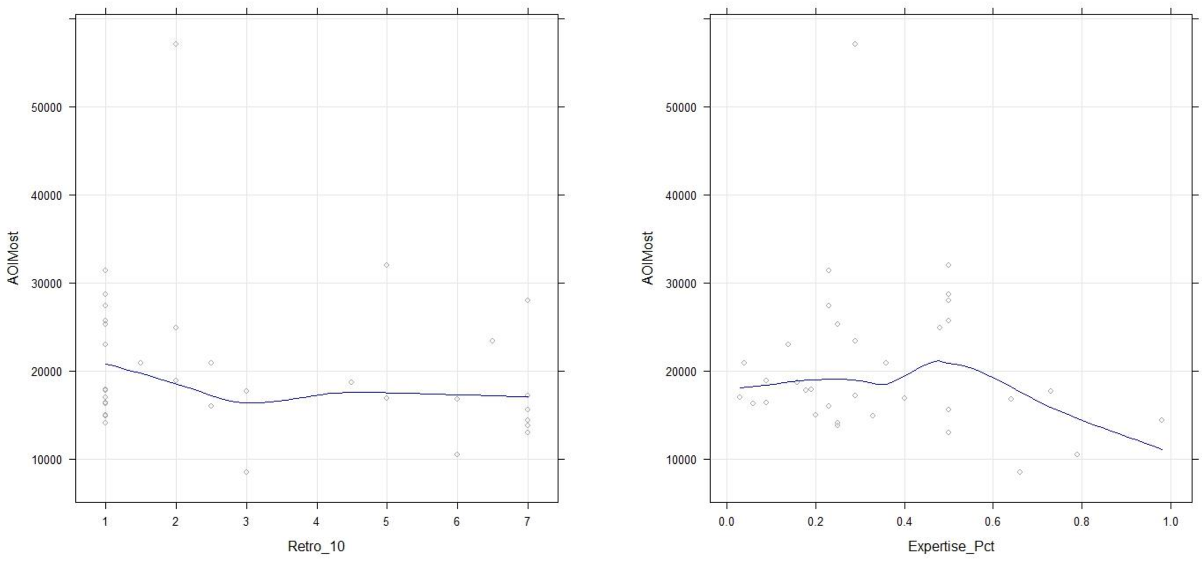


Figure 6.2: Processing related to exposure and self-rated expertise. Scatterplot to the left: Reading time (ms) of the *most* complex text (AOIMost) as a function of the exposure variable, use of Danish specialized texts (Exposure: Retro_10) measured on a seven-point Likert-scale. Scatterplot to the right: Reading time (ms) of the *most* complex text (AOIMost) as a function of self-rated expertise (Expertise_Pct).

The overall unexpected relatively short reading time of the most complex section might bias the results. It is likely that a fatigue effect is dominating the complexity of the least deep section placed after the deepest one (see figure C.4 of appendix C). It is also likely that complexity indicators of the two sections were too close. However, many other aspects may bias the performance (syntax, predictability and frequency or familiarity etc.) of the words to be read.

6.4.2 Correctness: Difficulty of content

As mentioned, I distinguish between the objective complexity indicators of a text and participants' subjective perception of the text difficulty, following Jensen (2009), but I would expect the perceived difficulty to depend on the complexity reflected in longer processing time of a complex text. In the retrospection, participants were asked to assess the difficulty of the chosen specialized text from the taxation discourse (see table 4.4). Participants evaluated the whole text, but we do not see any strong relation between subjective difficulty assessment and reading time (see figure 6.3).

Difficulty

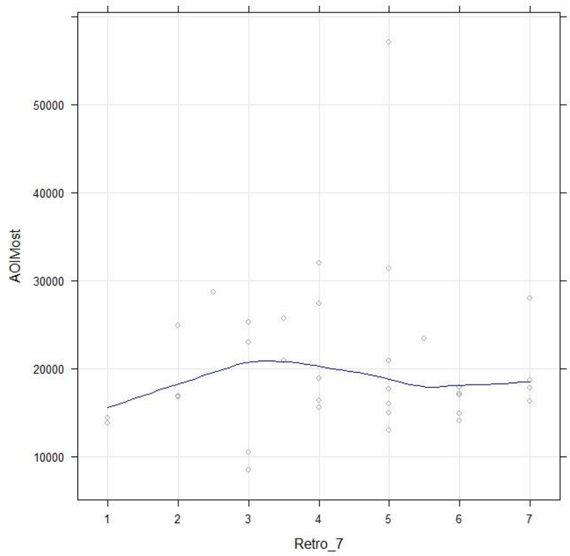


Figure 6.3: Processing related to peceived text difficulty. Scatterplot of the processing time (ms) of the *most* complex text (AOIMost) as a function of the perceived difficulty of the specialized text (Retro_7) measured on a seven-point Likert-scale.

6.5 Results: Knowledge structure

In the card-sorting task (see section 4.4.4), the highly subjective views are best reflected by depth compared to correctness or speed. The aim of the dissertation research is to further the interface design of a term bank, and I have chosen to emphasize the duality (textual versus graphical) information format of the dual-entry mode experiment. In particular, the graphical entry mode displays a concept diagram, which constitutes a terminological ontology complying with the formal structure of the domain's conceptualization.

Following Friege & Lind (2006), I begin with a quantitative evaluation of depth in a concept mapping test: Graph theory is used to measure the degree of interconnectedness (see table 6.10). I aim for the number of components i.e. the number of terms and relations. Friege & Lind (2006) uses the number of concepts connected with only one concept to reflect poor connectedness. However, in the domain many terms had only one relation (see table 6.9), and I therefore need to include a reference map (see figure F.2 in appendix F).

Table 6.9: Components of the reference map. The number of relations to other terms, for each of the 32 terms of the card-sorting exercise as they are drawn in the reference map (see figure F.2)

Number of Relations	Number of Terms
8	1
7	0
6	0
5	2
4	3
3	1
2	2
1	23
Total	**32**

Reference-map correspondence (see table 6.11) is a concept map's number of common edges with reference map. I use edit distance to measure the degree of correspondence. However, a reference map displays one of several views on the underlying conceptualizations, and therefore the evaluation may prove difficult, and I need to propose qualitative evaluation to infer expertise.

In the reference map, one term is related to eight other terms on the list, while 23 terms are only related to one other term on the list (see table 6.9). Consequently, the reference map contains 32 components. I chose to re-use the concepts displayed in the dual-entry modes for the structuring task. Ideally, the chosen concepts should avoid too many items with only one relation, and allow for more variance in the number of conceptual relations.

6.5.1 Correctness and depth: Evaluating concept maps

As mentioned, the quantiative evaluation was proposed by means of graph-theory components and reference-map correspondence.

Regarding the number of components, we realize that mean total components (50.6) is below the reference map (62.0) (see table 6.10). Some participants need more than the given list of 32 terms to draw their concept maps (ranging from 21 to 49 terms), while they are much more reluctant to draw relations (ranging from 5 to 46 relations), which means that all participants (apart from two) have fewer components than the reference map (see table 6.10).

Table 6.10: Interconnectedness of the produced diagrams. Interconnectedness of concept maps measured as the number of components, i.e. terms and relations included in the concept maps (disregarding the correspondence with reference maps).

Participant	Number of Terms	Number of Relations	Total components
1	32	32	64
2	40	9	49
3	36	6	42
4	49	46	95
5	27	25	52
6	28	5	33
7	21	11	32
8	26	22	48
9	23	17	40
Mean	**31.3**	**19.2**	**50.6**
Reference maps	32	30	62

Regarding the correspondence with the reference map, we realize that participants are capable of using a high number (mean 26.2) of terms from the given list of 32 terms in their drawings of concept maps (ranging from 18 to 32 terms), while the number of common edges (correct relations) are very small (ranging from 0 to 12 relations, with mean 5.3) compared to the reference maps containing 30 relations (see table 6.11).

Table 6.11: Reference-map correspondence. Correspondence between the reference maps and the produced concept maps measured as the number of common correct edges (relations) between terms.

Participant	Number of terms from the list	Number of correct Relations
1	32	10
2	32	0
3	30	2
4	30	9
5	27	12
6	25	0
7	21	2
8	21	10
9	18	3
Mean	**26.2**	**5.3**
Reference maps	32	30

The results of the quantitative evaluation leave an unclear result, which is difficult to relate to expertise. The problem is that high interconnectedness (many components) not necessarily reflects a high expertise because inconsistencies easily emerge as participants are eager to include as many terms as possible, while their structuring performance seem weak. On the other hand, I realize that a low reference-map correspondence is not necessarily reflecting low expertise, but rather a different view on the structuring. Therefore, I need a qualitative evaluation of the concept maps.

6.5.2 Inferring expertise from mapping strategies

If I approach the evaluation of the concept maps qualitatively, it becomes clear that participants' pursue different strategies in the card-sorting (structuring of terms) depending on their domain-specific expertise. I propose four types of mapping strategies: complex, listing, simplistic and absent (see table 6.12).

Table 6.12: Mapping strategies. Four mapping strategies are encountered with the nine participants (n=9) and ranked by level of complexity (complex, listing, simplistic, absent). The consistency of each map is ranked into three levels (high, medium, low). The strategy and consistency of the maps are used to infer levels of abstraction (researcher, practitioner, taxpayer) as indication of expertise.

Participant	Mapping Strategy	Structural Consistency	View on domain
1	Complex: deep	Medium	Researcher
4	Complex: broad	High	Taxpayer
5	Complex: augmented	Medium	Practitioner
2	Listing: 6 groups	High	Practitioner
3	Listing: 5 groups	Medium	Researcher
6	Listing: 4 groups	Medium	Taxpayer
8	Simplistic	Low	Taxpayer
7	Absent	Low	-
9	Absent	Low	-

Complex strategy

A complex strategy results in large diagrams containing many terms and relations and capturing different aspects of the domain as deep, broad or augmented with medium to high consistency (see table 6.12). Even though participants may have taken different approaches to producing the concept map (deep, broad or augmented), the complex concept maps produce a, quantitatively speaking, weak correspondence with the reference map, but they possess consistent structures (with few serious errors or misplacements) beginning with relevant overall subdivision criteria (direct versus indirect taxation or taxes versus duties).

As indicated in table 6.12, the three participants producing complex concept maps choose different strategies: Participant 1 chose a deep structure stressing the vertical connectedness (see figure F.3 in appendix F) covering exactly all the 32 terms, revealing an abstract “researcher” view to the task. Participant 5 chose the broad structure stressing the horizontal connectedness (see figure F.7 in appendix F) with a strong emphasis on the liability of personal taxpayers, which reveals a “taxpayer” view. Participant 4 produced an augmented structure (see figure F.6 in appendix F), where a number of added terms are necessary to complete the structure (almost like naming book shelves by theme and filling the terms inside). The structuring around the administrative systems collecting the taxes reveals a "practitioner" view. The deep structured concept map exhibit medium consistency, since the subordinates of corporate taxation is strangely misplaced and so is property taxation, while other parts of the structure are very precise, revealing that the participant knows his own line of business well, while his notion about the rest of the domain is vague.

Listing strategy

A listing strategy results in large concept maps with many terms but with very few relations constituting 4, 5 or 6 groups with medium or high consistency (see table 6.12). Three participants choose to draw concept maps looking very much like lists: Participant 2 structured the terms around the administrative systems collecting the taxes, which produced six groupings of high consistency and reveals a "practitioner" view (see figure F.4 in appendix F). Participant 6 used two groupings structured around the liability of taxpayers (personal versus corporate), which are added to the concept map below a direct and indirect taxation (see figure F.5 in appendix F). The focus on liability is revealing a "taxpayer" view, however the addition of the liability is highly misplaced and misleading, which means that collapsing that part would improve the consistency. Finally, participant 3 structured the terms into five groups based on a morphologic analysis of the contents, i.e. terms containing tax or duty, taxpayers, authorities, tax level, collection form (see figure F.8 in appendix F). The consistency seems high, but the problem is that seven terms are not placed or fitting into the proposed structure and that the first group contains the majority of the terms in the diagram (15 out of 25), which asks for a more consistent sub-division. The unconventional approach to the domain reveals a high level of abstraction in the structuring i.e. a "researcher" view.

Simplistic strategy

A simplistic strategy results in several small concept maps with relative few terms and relations with low consistency (see table 6.12). One participant chose the simplistic strategy: Participant 8 has produced a simplistic concept map beginning with the relevant overall subdivision criteria (direct versus indirect taxation), but that only covers 16 terms, while another diagram contains 8 terms in an absent structure (see figure F.10 in appendix F). The low coverage and weak consistency suggests a "taxpayer" view.

Absent strategy

An absent strategy results in small concept maps with few terms and few or absent relations with low (absent) consistency (see table 5.4). Participants 7 and 9 have attempted to produce concept maps showing the relations between the listed terms, but there is no clear strategy or structure, and the two participants begin with a particular tax type: central government tax and corporate tax (see figure F.9 and F.11 in appendix F). Moreover, the terms are strangely misplaced, repeated or without relation to other terms in the concept map. The absent strategy and consistency does not point to any particular view.

6.6 Conclusion

I conclude that using card-sorting exercises as indirect measures of expertise prove very difficult, because the descriptive evaluations are complicated, which challenges the inference of expertise. However, the concept maps constitute beneficial input from the target users to the graphical design of the term-bank interface.

In three representative tasks relevant to domain-specific terminology, I proposed associated expert performance indicators:

Firstly, we conclude that in an open-ended, association task of recalling taxation terms experts listed more terms exhibit higher depth (the response time was not recorded). Secondly, I conclude that in a categorizing taxation terms and pseudo-terms, experts with high exposure and motivation are more correct and offer definitions of more precision and detail (again response time was not recorded). Thirdly, in a reading task using an authentic, specialized text containing long and complex, domain-specific terms expert performance is not faster compared to novice performers, but experts perceive the contents as less difficult (due to higher depth). Fourthly, I conclude that inferring expertise from the quantitative evaluations of the concept maps by means of interconnectedness or reference-map correspondence are not appropriate and concept maps are highly subjective. Therefore, I propose a qualitative evaluation and conclude that four mapping strategies comprising complex, listing, simplistic and absent strategies appear with three levels of abstraction in the views on the domain (researcher, practitioner and taxpayer).

The weak expertise effects of the representative tasks have three possible interpretations: It is possible that expertise effects are hard to demonstrate in the taxation domain, as we are all exposed to taxation terminology due to our tax liabilities. It is also possible that the representative tasks (dependent variables) do not fully reflect expertise characteristics (correctness, fastness and depth) and need a further development on all three task designs to improve the number of observations and to better capture the associated indicators of superior expert performance. Finally, it may be that the (explanatory) expertise variables are biased and need further improvement (see section 5.5).

Regarding the research theme, I conclude that there is a need for further development of indirect measures of expertise, before we may expect them to be significant in explaining the variance of target users' dual-entry performance in a regression approach (see chapter 7).

Chapter 7: Dual-entry mode experiment

Chapter 7 contributes to analyzing the final underlying research theme linking target-user performance to expertise and learning. I show that overall domain-specific terminology and knowledge in the taxation domain can be transferred to target users across different levels of expertise (despite weak expertise effects) and by means of dual complementary information entry modes (concept diagrams and concept articles).

In this chapter, I review and combine current knowledge of eye-tracking, knowledge acquisition and multimodal information search (section 7.1). The regression approach is described and the selection of dependent and explanatory variables is motivated (section 7.2). The results of the correctness model are presented and discussed (section 7.3), which is refined into a response-time model (speed) (section 7.4) and a diagram-fixation model (depth) (section 7.5).

7.1 Introduction

De Schryver (2003) systematically outlines an extensive list of dreams of the lexicographer compiling electronic dictionaries, and the majority of dreams pertain to the data access of target users. In particular, the electronic dictionary renders possible the multimodal representation of meaning to its target users (Lew, 2010). An important distinction between different motivations for information search is whether target users apply random or deliberate searching strategies. Serendipity (discovery by accident) is likely to cause inefficient processing times due to long scan paths (see section 4.1 about eye tracking), but it is not necessarily a problem if random search strategies are underlying information retrieval. Target users may even show a preference for applying a browsing strategy which provides answers merely by chance. However, we should strive for an experimental design where the chances of retrieving the answers by chance are minimized, otherwise we are not accurately attribute an observation to specific causes (Usability First, 2015). In addition, we should avoid overloading the limited cognitive capacity. We may interpret the term bank as type of instructional design enabling and facilitating learning and knowledge acquisition.

The total cognitive load is the sum of the intrinsic load (ability to represent task information effectively), the extraneous load (presentation of task) and germane load (organize knowledge structures), and if the load exceeds the capacity, learning is impaired as suggested by Appel & Kronberger (2012). In the process of learning, animations may support both an enabling function (allow for cognitive processing otherwise impossible) and a facilitating function (allow for cognitive processing otherwise demanding high mental effort). Schnotz & Rasch (2005) compares to the cognitive load theory, where facilitation resembles redundancy, which increases the extraneous cognitive load (redundancy) due to the processing of unneeded information. However, it may also be interpreted as unintended decrease in germane cognitive load (mental capacity), because unused mental capacity impedes learning processes.

Education researchers investigate the conditions, where learners benefit the most from multimedia learning materials, which are highly relevant to modern term bank users searching multimodal user interfaces. Brünken, Plass & Leutner (2004) examine the limited cognitive capacities of different subsystems of the working memory by combining visual presentations (textual and pictorial material) with audiovisual presentations (narrations and pictures) and auditory (music or sounds). Brünken, Plass & Leutner (2004) discuss an attentional adaptation effect, which is the result of the meta-cognition of participants in longer experiments allowing participants to ignore irrelevant information (e.g. auditory). Cuevas, Fiore & Oser (2002) investigate learning (meta-cognitive) and the processing of text (verbally) and diagram (visually), which is activating different mechanisms reinforcing encoding of participants.

Mayer & Moreno (2010) investigate multimedia instruction presenting words (e.g. on-screen text) and static pictures (e.g. graphs or maps) intended to foster multimedia learning i.e. to integrate presented material with existing knowledge. In particular, the dual channel assumption of the information processing system is key to the cognitive load reduction proposals. Pacharapha & Ractham (2012) study motivational factors behind knowledge transfer, which involves two parties i.e. the source and receiver of knowledge. On the source side, knowledge transfer requires the willingness to share knowledge. On the recipient side, knowledge acquisition requires learning. Goldberg, Stimson & Lewenstein (2002) underlines that eye tracking on the search of user interfaces is necessarily constrained to specific stimuli.

I emphasize knowledge acquisition, i.e. (human) target users acquiring knowledge from the dual-entry modes. In the experiment, participants were presented with static images resembling complementary term bank entries in text and graphics. It was a highly controlled experiment conducted in an eye-tracking laboratory, which was meant to be a contribution to an interface design process of a term bank containing domain-specific terminology.

7.2 Regression approach

Two complementary information modes (text and graphics) allow users to access the data of the term bank from dual access points. The images are static and do not allow for any dynamic interaction. At this experimental stage the purpose is to further the interface design process, i.e. the results will inform the re-design of the interface, which is then to be re-tested, until a satisfactory usability level is achieved.

Significant predictors

For the analysis, a regression approach is applied. Multiple regression techniques allow for the assessment of multiple correlations of explanatory (independent) variables with the dependent variable (Balling, 2008, p. 94). Hence, a regression analysis makes it possible to determine whether there are effects (i.e. significant predictors) of each explanatory variable dominating the other explanatory variables included in the regression model. In addition, the approach allows for statistical control (as opposed to experimental control), i.e. allows us to isolate marginal effects given all the other variables. Finally, the regression approach allows for the inclusion of both numerical variables, e.g. self-rated expertise and exposure, and categorical variables, such as the question type category with three different levels (D, A and DA). In particular, we avoid dichotomisation (i.e. variables only exhibiting oposing values) and consequent loss of statistical power.

In the regression design, several statistical choices exist for the data analysis, however, linear mixed-effects modelling seems to be the most powerful without being anti-conservative, i.e. likely to result in a so-called type-1-error i.e. (falsely) rejecting a null hypothesis which is in fact correct. In particular, mixed-effects models include both fixed (repeatable) and random (non-repeatable) variables, the latter including so-called random intercepts (reflecting participant level) and random slopes (reflecting participant profiles), allowing us to assess whether group differences are significant over and above differences between individual participants. It is a method that allows us to model dependencies in the observations, e.g. the answers of each participant are not considered independent. This means that we may infer learning and expertise effects from the regression models.

Eye movements

In eye-tracking research, the recorded eye movements are analyzed by means of detecting events, i.e. measures accounting for scan paths and fixation duration (Holmquist et al., 2011). In particular, the eye-mind-hypothesis (Just and Carpenter, 1980) ascertains that fixations on a stimulus indicate the cognitive effort needed to process and understand that stimulus. The response time comprises mainly the sum of fixation and scan paths, but the experimental design also allows participants to look elsewhere, i.e. outside the screen where eye movements are recorded. This is e.g. the case when participants look down to key in their answer.

Performance models

I use the eye-tracking observations to examine participants' performance, i.e. their dual-entry mode processing associated with answering each of the multiple-choice questions. In particular, I use the regression approach to investigate expertise effects on partipants' performance (i.e. whether expertise variables are significant predictors of performance), and that requires a suitable dependent variable, which reflects the characteristics of expert performance. In chapter 6, the proposed performance indicators were reflecting characteristics of expert performance (correctness, speed and depth), and they were used to evaluate performance in each of the representative tasks comprising recalling, categorizing, reading and structuring terminology. The representative tasks developed in chapter 6 are not fully compatible with the task of the dual-entry mode experiment, but the performance-indicator framework is applied in the selection of dependent variables in the regression approach.

I investigate three properties important to the usability of a formal (displaying terminological ontologies) and a (user-oriented) functional term bank: Firstly, a term bank should allow users to acquire knowledge about a domain, this property will be evaluated by analyzing users' learning outcome, indexed by success in the multiple-choice questions. Secondly, a term bank should facilitate easy access to the data of the term bank avoiding any redundancies or overload, which will be evaluated by analyzing response times. Thirdly, a term bank should include a graphic entry mode complementary to the traditional written entry mode, which will be evaluated by participants' use of concept diagrams measured as their diagram-fixation time on correct answers. The performance indicators used to evaluate the representative tasks outlined in chapter 6 are applied as dependent variables in three regression analyses (performance models) of the dual-entry modes.

Correctness is the first performance indicator and dependent variable of the first regression analysis, which I call the correctness model (see section 7.3). Then the second performance indicator (speed) is included by using response time on correct answers as a dependent variable, the so-called response-time model (see section 7.4). Finally, the third performance indicator (depth) is included by using fixation time on diagrams on correct answers as dependent variable, the so-called diagram-fixation model (see section 7.5).

The dependent variables are used in linear mixed-effects regression models with question, participant and block as random effects, i.e. answers to each of the 48 questions in each of the 8 blocks are not considered independent. Linear mixed-effects models are available in the packages lme4 (Bates, Maechler, Bolker & Walker, 2013) and languageR (Baayen, 2011) within the statistical computing environment R (mainly version 3.0.2, R Core Team, 2013). These models allow for the modelling of non-linear as well as linear effects. A bottom-up approach was used, testing variables one at a time, starting with the most control-oriented and ending with those most central to the hypotheses (see table 7.1). Only significant variables were retained in the final analysis reported below.

Table 7.1: Overview of regression models. Performance indicators are dependent variable in each of the performance models. "SIG" reflects a significant effect, while "NS" indicates non-significant effect, and grey colour indicates that the variable was irrelevant for that model. The full regression models are outlined in the tables stated in the parentheses. "Q" indicates a non-linear effect, "POS" indicate a positive effect and "NEG" a negative effect. Explanatory variables are ordered by importance beginning with the least important.

	Dependent variables		
	Correctness	**Speed (Response time)**	**Depth (Diagram fixation)**
Explanatory variables	**(see table 7.3)**	**(see table 7.4)**	**(see table 7.5)**
Display side of answer	NS	NS	SIG (NEG)
Total response time	SIG (NEG)		
Self-rated search expertise	NS	NS	NS
Number of weekly Google search	NS	NS	NS
View on A mode	NS	NS	NS
View on D mode	NS	NS	NS
View on performance in A	NS	NS	NS
View on performance in D	SIG (POS)	NS	NS
Preference for D compared to A	NS	NS	NS
Preference for None compared to A	NS	NS	NS
View on information modes	SIG (POS)	NS	NS
Self-rated tax expertise	NS	NS	NS
Exposure to specialized texts	NS	NS	NS
Motivation	NS	NS	NS
Age	NS	NS	NS
Gender	NS	NS	NS
Question type D compared to A	NS	SIG (POS)	SIG (POS)
Question type DA compared to A	NS	NS	SIG (POS)
Trial number	SIG (POS)	SIG (Q)	SIG (Q)
Block trial number	NS	SIG (NEG)	NS

Only a limited number of the variables in each of the performance models turned out significant. However, trial number is significant in each of the models, indicating that participants’ performance evolves as the experiment proceeds (local learning effects). Below we present and interpret each of the performance models. I expect expertise effects on participants' processing of the dual-entry modes and answering multiple choice questions. Significant expertise effects have the potential of guiding the user adaption of the entry modes of the term bank, in particular, the dissemination of knowledge to professional target users with different levels of expertise, in general.

7.3 Results: Correctness

I investigate the outcome of the experiment measured as correctness on the multiple-choice questions. A-questions produced the most correct responses (76.6%), DA-questions the second-most (65.8%) and D-questions produced the lowest number of correct responses (57.6%). The data of three participants were excluded from the analysis because the eye-tracking system failed to record during two of the sessions, and one participant misunderstood the instructions and considered the entry modes as distractors, which she ignored, and that resulted in a very high error rate. One outlier was excluded from the analysis, so the analyses reported are based on data from 36 participants. A summary of the regression model for correctness is presented in table 7.2.

The estimated coefficients for the different variables (continuous or factor levels) are summarised in the second column (denoted “Estimate”) and the associated p-value (based on the t-distribution) in the other column (denoted “p(t)”). The “Intercept” is the value of the dependent variable in the (hypothetical) case where all predictors are zero.

Table 7.2: Summary of regression model for correctness. Random effects for 48 questions, 36 participants and 8 blocks. Reference level for correctness is correct. The self-assessed performance on diagrams (Retro_6) and the view on information coverage (Retro_2) is measured on seven-point Likert-scales.

FIXED EFFECTS

	Estimate	p(t)
Intercept	5.4386	0.0025
Log (Response time)	-0.7003	<0.0001
Self-rated performance (Retro_6)	0.2375	0.0331
Information coverage (Retro_2)	0.3816	0.0072
TrialNo	0.0205	0.0044

RANDOM EFFECTS

	Std. Dev.
Question	0.7966
Participant	0.5134
Block	0.5837

The regression model shows four significant predictors (explanatory variables) for the dependent variable correctness. The signs of the estimated coefficients of the significant predictors are interpreted (section 7.3.1) and the results are related to existing literature (section 7.3.2).

7.3.1 Significant predictors of correctness

In figure 7.1, the partial effects of the siginificant predictors in the correctness model are shown, i.e. response time, self-assessed performance on diagrams, view on information coverage and trial number. The figure shows the effect of a given predictor with all other predictors held constant at the median values for continuous predictors.

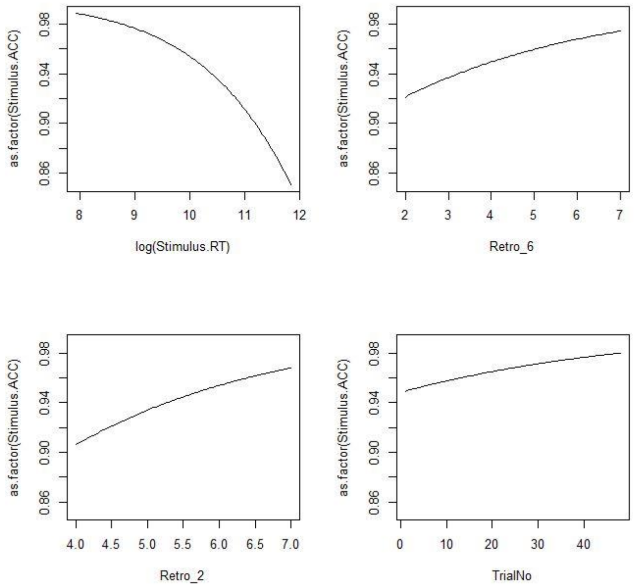


Figure 7.1: Predictors of correctness. Partial effects of the four significant predictors in the regression with correctness as dependent variable. The vertical axes show the probability of correctness. Top left: Response time (log(Stimulus.RT)). Top right: Self-assessed performance on diagrams (Retro_6, measured on a seven-point Likert-scale). Bottom left: Assessment of information coverage (Retro_2, measured on a seven-point Likert-scale). Bottom right: Trial number (TrialNo).

(1) Response time is a significant predictor of correctness (p-value < 0.0001) such that faster response times resulted in more correct answers. It should be noted that the response time variable is logarithmically transformed to reduce the skewness which might otherwise distort the results. Response time has a significant effect on correctness, and the estimated effect is negative i.e. the longer response time of the participant, the lower correctness, and more errors. In other words, participants spending a long time searching the entry modes for an answer will get fewer correct answers, reflecting either that the participant is not capable of understanding the question because it is difficult or confused by the available answers or the entry mode providing the answer. Despite this seemingly counter-intuitive result that increased effort decreases success in the experiment, correctness is high overall. In particular, a learning effect emerges, increasing the participants' understanding and success as the experiment proceeds (see the trial-number effect below).

(2) Self-assessed performance on diagrams (denoted Retro_6) ranging from 1 to 7 is significant (p-value 0.0331), with the exact wording of the question: "How well do you think that you performed when you retrieved answers in diagrams? Please, qualify" (see table 4.4). The participants' self-assessment of their performance on diagrams has a positive correlation with correctness, i.e. the better participants assess their performance on the questions where they found the answer in the diagram entry mode, the higher the correctness. The concept diagram is the least familiar entry mode to most participants, whereas the concept articles more closely resemble well-known written information templates, raising participants' self-assessed performance, or confidence, on concept articles (75.0% of answers being more than or equal to 6) compared to diagrams (61.1% of answers being more than or equal to 6). If the D-questions are more error-prone (producing relatively low correctness) in the experiment, where the overall correctness is relatively high, and the A-questions cause little difficulty and show a non-significant effect on correctness. It is likely that participants, who encounter few problems with diagrams, thus rating their performance well, do actually show a high level of correctness. The upward slope in the top right panel of figure 7.1 reflects the facilitatory effect on correctness, i.e. higher self-assessed diagram performance increases correctness.

(3) View on information coverage (denoted Retro_2) ranging from 1 to 7 is significant (p-value 0.0072) with the wording of the question: "And how well do you think you got your information need covered (i.e. did the articles and/or diagrams provide you with an answer to the posed questions? Please, qualify" (see table 4.4). This effect validates the self-assessment measure of participants' assessment of information coverage in the retrospective interview, because participants' perception of the information coverage has a significant effect on the actual correctness, i.e. outcome, of the experiment. The upward slope in the bottom left panel of figure 7.1 reflects the facilitatory effect on correctness, i.e. higher assessment of information coverage reflects equally high performance in terms of correctness.

(4) Trial number ranging from question number 1 to 48 (see appendix D) is significant (p-value 0.0044) in accounting for the variance in correctness. The estimated coefficient of trial number is positive i.e. the higher trial number, the

higher correctness, reflecting a learning effect. This increase in correctness indicates an increase in their understanding of the questions and entry modes as the experiment proceeds, which we may interpret as a learning effect. The upward slope in the bottom right panel of figure 7.1 reflects the facilitatory effect on correctness, i.e. higher trial number increases correctness.

Interestingly, the analysis shows no expertise effects. Neither the expertise variables, participation (work place), motivation (work area), education (length), exposure to the specialized discourse (taxation), nor the self-rated expertise ratings are significant in explaining correctness. However, trial-number is significant, which we may interpret as a learning effect, since participants increase performance (correct answers), i.e. acquire knowledge, as the experiment proceeds.

7.3.2 Discussion of correctness

Participants were instructed to invest the time needed to find the correct answer rather than performing fast which would potentially have resulted in high error rates. It therefore comes as a surprise that participants spending a long time before answering the questions are more likely to get it wrong. The reason lies either with the multiple-choice question format including the available answers or with the dual-entry modes containing the answer. I refine the correctness analysis in a regression model with the response time on the correct answers as the dependent variable in section 7.4.

Following Rikers & Paas (2005), we expect experts to perform representative tasks (almost) error-free. I did not encounter any expertise effect in the correctness model above (see table 7.1), as self-rated expertise has a non-significant effect on correctness (p-value 0.4971). There are several reasons why an expertise effect may be reduced, absent or reversed:

In the case of reduced expertise effects, experts perform marginally above non-experts: The dual-entry mode stimulus-pairs challenge, and potentially overload, the limited human cognitive capacity (Mayer & Moreno, 2003; Brünken, Plass & Leutner, 2004 and Appel & Kronberger, 2012) available for information processing.

In the case of absent expertise effects, experts are not performing different from non-experts: the direct measures (chapter 5), as well as the indirect measures (chapter 6) of domain-specific expertise, may be inadequate due to the complexity and inter-disciplinarity of the social sciences (Alexander, 1992) comprising the taxation domain. Therefore the (one-dimensional) domain-specific expertise measures/variables do not exactly reflect the (multi-dimensional) domain-specific expertise and do not significantly account for the variance in correctness (or response time or diagram-fixation time, analysed below). It could also be the case that the performance of the non-adapted dual-entry modes actually triggers learning effects as the experiment proceeds overriding potential expertise effects.

In the case of reversed expertise effects, experts perform under non-experts: The performance on the multiple-choice-question format may be favoured (general) information search skills rather than (domain-specific) expertise. In particular, the experiment may open for serendipitous information encounters (Foster & Ford, 2003), allowing participants to retrieve answers accidentally by browsing in no particular systematic way. In addition, expertise performance is not facilitated by the dual-entry modes, which requires non-domain specific (short-term) working memory instead of (long-term) domain-specific (declarative and procedural) knowledge (Alexander, 1992, Kuhn, 2000, and Mayer & Moreno, 2003). Finally, experts apply stereotypical and inflexible problem-solving strategies to tasks, which are uncommon to them (Bilalić, McLeod & Gobet, 2008).

Under the expertise reversal effects (Kalyuga & Sweller, 2004), the expert performance is inhibited by a redundancy effect providing participants with information or explanations they already possessed. Hence, it could very well be the case that experts had performed better in a test only comprising the multiple-choice questions without the dual-entry modes. However, with the dual-entry modes displayed, participants spend time processing them, indicating an enabling function (making answering possible) rather than the intended facilitating function (making answering effortless) (Schnotz & Rasch, 2005). A few participants did voice their surprise when they caught themselves looking for answers, they "already knew."

I expected high correctness to be an indicator of expert performance. However, none of the expertise variables are significant predictors of correctness, i.e. expertise effects are eliminated, presumably due to a combination of the aspects, mentioned above. If the usability of the term bank is merely a question of efficacy (outcome), no user adaption according to expertise seems to be necessary, rather the users train themselves to use the term bank (learning effect).

7.4 Results: Response time

Response time is defined as the sum of participants' processing and answering time on each question. Participants control the speed of the experiment, and proceed to the next question and stimulus-pair by pressing the space bar, which indicates the start of response time for a given question. The end is indicated when participants answer the question by hitting 1, 2 or 3 on the keyboard.

Correctness turned out very high, which means that participants overall were able to understand the questions, acquire knowledge from the entry modes and find the correct answers, and despite a negative response-time effect, they showed increasing correctness as the experiment proceeds (positive trial-number effect in the correctness analysis). Besides answering correctly, a relevant performance indicator is to consider the response time (speed) needed to produce the correct answers. Response time includes participants' total processing time: reading the multiple-choice question, finding the answer in one of the entry modes, and the time spent choosing the correct answer, i.e. finding and pressing the correct number on the keyboard. A summary of the regression model for response time on correct answers is presented in table 7.3.

Table 7.3: Summary of regression model for response time on correct answers. Reference level is question type article. Random effects for the 48 questions, 37 participants and 8 blocks. “Name” specifies whether the standard deviation refers to random intercepts or to random slopes for a particular variable, in this case trial number. The intercept (9.9940) is log-transformed.

FIXED EFFECTS

	Estimate	p(t)
Intercept	9.9940	N.a.
QuestionTypeD	0.3234	0.0132
QuestionTypeDA	-0.1106	0.3805
TrialNo	-0.0385	<0.0001
TrialNo^2	0.0005	<0.0001
BlockTrialNo	-0.0013	0.0276
<u>Interaction</u>		
QuestionTypeD:TrialNo	-0.0001	0.9460
QuestionTypeDA:TrialNo	0.0042	0.0159

RANDOM EFFECTS

	Name	Std. Dev.
Question	Intercept	0.3240
Participant	TrialNo	0.0070
	TrialNo^2	0.0001
Block	Intercept	0.5349

The regression model for response time shows three significant predictors and one significant interaction for the dependent variable response time. Random slopes are included for trial number by participants, corresponding to (slightly) different trial number effects for each participant. The signs of the estimated coefficients of the significant predictors are interpreted (section 7.4.1) and the results are related to existing literature (section 7.4.2).

7.4.1 Significant predictors of response time

In figure 7.2, the partial effects of the significant predictors of the response-time model are shown, i.e. trial number effects inside the block and across the entire experiment, the latter varying between the three question types.

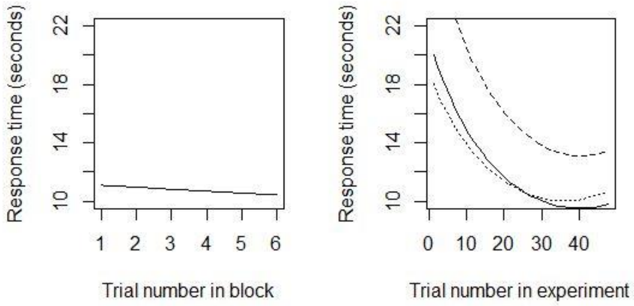


Figure 7.2: Predictors of response time on correct answers. Partial effects of the two significant predictors in the regression model with response time on correct observations as dependent variable. The vertical axes show response time in seconds. Left: Trial number effect in block (BlockTrialNo). Right: Trial number effect in experiment (TrialNo). The three lines represent the question types (dashed line is Diagram, solid line is Article, and dotted line is Diagram-Article).

(1) Block-trial number ranging from 1 to 8, where each of the blocks represents a target term and a fixed entry-mode stimulus-pair, to which six randomized questions belong, has a significant effect (p-value 0.0276) on the response time on correct answers. The estimated coefficient of the block-trial number effect is negative (confirming the overall trial number effect in the correctness model), i.e. the more questions answered inside the block, the shorter (log) response time. The (weak) downward slope in the top left panel of figure 7.2 reflects the facilitatory effect, but the effect is not very strong. Participants familiarize themselves with the dual-entry mode as the questions of the block are being answered.

(2) Question type has three conditions (A denotes article, D denotes diagram and DA denotes diagram-article based question types) with the reference level being A-question. Overall the question type is significant in accounting for variance in the dependent variable, response time on correct answers. The estimated coefficient of D-questions is positive indicating that D-questions require significantly longer response times (compared to A-questions, p-value 0.0132). This is illustrated from the systematically higher (dashed) curve in the right panel of figure 7.2. The corresponding difference between DA-questions and A-questions is non-significant (p-value 0.3805). One possible interpretation is that the participants are choosing the presumably more familiar (and less time-

consuming) article entry (or avoiding the less familiar and more time-consuming diagram entry) to answer the DA-questions, because we see a significantly higher response time for D-questions compared to both A- and DA-questions (see the higher position of the D-curve in the right panel of figure 7.2). However, it is also possible that participants are able to use the diagram entry faster in answering the DA-question compared to D-question, because they may find the DA-questions easier. It is necessary to consider the eye-tracking data to determine participants' actual use of the dual-entry modes (see section 7.5).

(3) Trial number shows a (non-linear) significant effect (p-value <0.0001) on response time (see right panel of figure 7.2), which we may interpret as an overall learning effect, but the slope (learning) is steeper for low trial numbers, and flattening out as the trial number increases. However, the model shows an interaction between the explanatory variables trial number and question type, with a significant difference in the shape of the trial effect (p-value 0.0159) between the conditions DA and A, but non-significant difference (p-value 0.9460) between the conditions D and A. Hence, the trial-number effect for DA-questions differs significantly from the A- and D-questions (see the different profile of the DA-curve in the right panel of figure 7.2), reflecting a different learning effect on the DA-question, where the answer can be found in both entry modes. It should be noted that the increase in the three curves of figure 7.2 towards the end of the experiment not necessarily reflect decreased learning. In both ends of the curve, the development is based on few observations and therefore not very reliable.

The response-time model shows no expertise effects (in line with the correctness model), but there seem to be differences in the trial-number effect across participants (see figure 7.3), these are modelled by including so-called random slopes in the regression analysis.

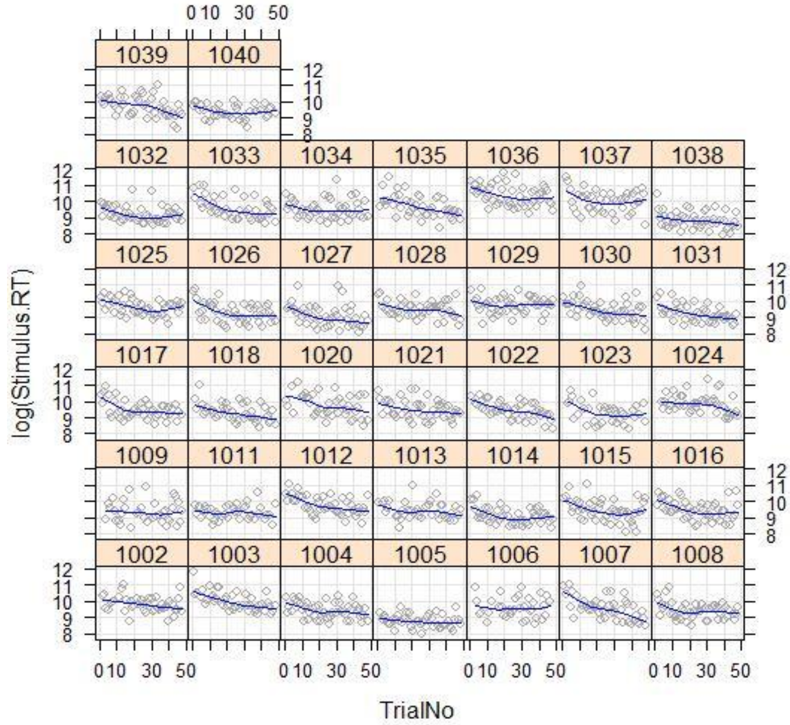


Figure 7.3: Random slopes of each participant. Scatterplots of raw data for the potential random slopes for participants.

Including random slopes for participants in the regression model confirms the overall significant trial-number effects, including the significant differences between DA- and A-question types, but non-significant differences between D- versus A-question types. If we interpret trial number effects as learning effects, the regression model for response time shows some differences in participants' learning profiles, as illustrated in figure 7.3, but nonetheless an overall pattern of different learning profiles for different question types, as shown in figure 7.2.

To sum up, the response time for D-questions is longer than A- and DA-questions, which means two things: it takes longer to (correctly) encode the graphics compared to text, and participants prefer to use the quicker article entry to find the answer, when it is possible (i.e. in A- and DA-questions). Overall, learning effect in the response-time model (speed) on correct answers is maintained (compared to the correctness model) reflecting that participants increase performance (speed on correct answers) as the experiment proceeds.

7.4.2 Discussion of response time

Following Rikers & Paas (2005), we would expect experts to perform representative tasks faster. As in the correctness model, there were no significant expertise effects in the response-time model (see table 7.3), as self-rated expertise (as well as the other expertise variables) showed a non-significant effect on correctness (p-value 0.6780). We may repeat the reasons, why an expertise effect may be reduced, absent or even reversed (see section 7.3.2).

I expected speed and correctness to be indicators of expert performance. However, as we saw in the simple correctness model (see the discussion of reduced, absent and reversed expertise effects in section 7.3.2), neither of the expertise variables nor the self-rated domain-specific expertise are significant predictors of response time on correct answers. If the usability of the term bank is merely a question of efficacy (success) and efficiency (success and response time), no user adaption regarding domain-specific expertise is relevant. It seemed from the illustration of random slopes that participants' learning profiles would differ (see figure 7.3), which might have proposed a case for user adaption for something different than expertise.

7.5 Results: Diagram fixation

In the dual-entry mode, the concept-oriented diagram is complementing the term-oriented articles expanding the amount of presented knowledge to the target user. In particular, the diagram constitutes the novel aspect of the entry-mode and target-user processing, and usability of the diagrams will determine whether the proposed dual-entry mode should be maintained or rejected.

As mentioned, three so-called areas-of-interest (AOIs) are introduced, one for the question area and one for each of the two types of entry mode (see figure D.2 in appendix D). Gaze duration is measured for each of the three AOIs, defined as the sum of all fixations over 200 ms on the screen in the relevant area. We begin by excluding observations showing question-fixation time equal to zero, because something must be wrong with the eye-tracking measurement, since participants necessarily must fixate on the question to be able to answer, while the same does not hold for fixations in the diagram or article AOIs, which meaningfully can take the value zero.

In the response-time model, we saw that participants spend longer time answering D-questions correctly compared to the A-questions (and DA-questions) (see figure 7.2). The eye-tracking data will contribute to our understanding of participants' use of the dual-entry mode, i.e. whether participants are using the diagram to answer the DA-question. A summary of the regression model for diagram fixation on correct answers is presented in table 7.4.

Table 7.4: Summary of regression model for diagram-fixation time on correct answers. Reference level is question type article. Random effects for the 48 questions, 37 participants and 8 blocks.

FIXED EFFECTS

	Estimate	p(t)
Intercept	27.86	<0.0001
DisplaySideDiagram_Right	-6.14	<0.0001
QuestionTypeD	48.15	<0.0001
QuestionTypeDA	22.74	<0.0001
TrialNo	-153.93	<0.0001
TrialNo^2	153.11	<0.0001
QuestionFix	453.20	<0.0001
QuestionFix^2	-61.89	0.0044
<u>Interaction</u>		
QuestionTypeD:TrialNo	4.62	0.9264
QuestionTypeD:TrialNo^2	-139.77	0.0045
QuestionTypeDA:TrialNo	100.48	0.0418
QuestionTypeDA:TrialNo^2	-53.64	0.2698

RANDOM EFFECTS

	Std. Dev.
Question	6.03
Participant	6.71
Block	0.00

The diagram-fixation model shows four significant predictors including one significant interaction for the dependent variable diagram fixation time on correct answers. The signs of the estimated coefficients of the significant predictors are interpreted (section 7.5.1) and then the results are related to existing literature (section 7.5.2).

7.5.1 Significant predictors of diagram fixation

In figure 7.4, the partial effects of the significant predictors in the diagram-fixation model are shown, i.e. display side of diagrams, question-fixation time and trial number.

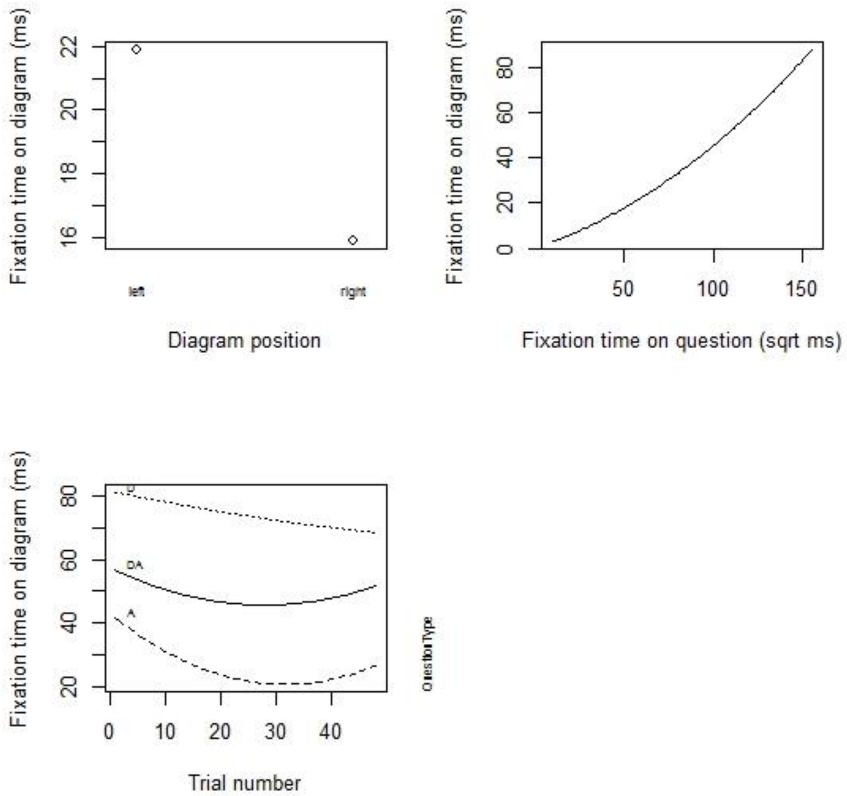


Figure 7.4: Predictors of diagram fixation on correct answers. Partial effects of the three significant predictors in the regression model with diagram fixation on correct observations as dependent variable. The vertical axes show fixation time on diagrams in ms. Top left: Diagram position. Top Right: Question-fixation time. Bottom Left: Trial number effect in experiment (TrialNo). The three lines represent the question types (dashed line is Diagram, solid line is Article, and dotted line is Diagram-Article).

(1) Display side of diagrams takes two levels, left or right, which is randomized by *E-prime*. Diagrams displayed to the right have a negative significant effect (p-value < 0.0001) on diagram-fixation time on correct answers, i.e. diagrams displayed to the right (left) imply shorter (longer) diagram-fixation time. A possible interpretation is that participants have a preference for searching (and reading) from the left to the right, i.e. entry modes displayed to the left are processed first and therefore slightly longer, especially in the cases where the participant has to realize that the answer lies in the complementary entry mode.

(2) Question-fixation time is the total fixation time (gaze duration) inside the question AOI placed in a horizontal field above the stimulus-pair (see figure D.2 in appendix D). Fixation time is recorded by the eye tracker and shows a (non-linear) significant (p-value 0.0044) effect on diagram-fixation time. The upward slope of the graph in the top right of figure 7.4 indicates that long question fixation is associated with increases in diagram fixation time. It probably reflects a longer fixation on the more difficult questions, which is distributed across the AOIs.

(3) Question types D and also DA show significant differences (p-values < 0.0001) compared to A-questions on diagram-fixation time. (In the response-time model, DA-questions did not differ significantly from the A-questions). As indicated by the position of the graphs in figure 7.4 (bottom left panel), the diagram-fixation time is longest on the D-questions, second-longest on the DA-questions, and shortest on the A-question types. It is not possible to determine in which entry mode participants actually find the answer. In principle, the last fixation point would be a good indicator, but in practice the dual-entry modes are maintained inside the block and the answer might therefore have been retrieved earlier, or be the result of a random search of both entries.

In eye-tracking research, the eye-mind-hypothesis (Just & Carpenter, 1980) implies that fixations on a stimulus can be interpreted as the cognitive effort needed to process a stimulus. My experimental design allows me to evaluate whether fixations are relevant in a given question type, which opens for three interpretations (see table 7.5):

Table 7.5: Interpretation of diagram-fixation time. Diagram fixation is ranked (high, medium, low) according to figure 7.4 (bottom-left panel) on each question type. The importance of diagram fixation for answering each question type is stated (indispensable, voluntary, irrelevant). The interpretation of the diagram fixation constitute (demanding, effortless and distracting) diagram use.

	Diagram fixation	**Question type**	**Diagram importance**	**Interpretation of diagram use**
(i)	High	D-question	Indispensable	Difficulty
(ii)	Medium	DA-question	Voluntary	Effortless
(iii)	Low	A-question	Irrelevant	Distraction

Ad (i) Diagram-fixation in the D-questions is presumably indispensable for finding correct answers, which possibly reinforces the diagram-fixation over and above the other two question types. In this case, diagram-fixation reflects the demanding or cognitive effort that participants experience. Ad (ii) Diagram-fixation in the DA-questions is voluntary for finding the answer, which also appears in the article entry mode. In addition, we cannot be sure that the answer is found in the diagram, but we do know that the diagram-fixation time is below D-questions. In this case, diagram fixation must necessarily reflect an effortless diagram use rather than demanding diagram difficulty. Ad (iii) Diagram-fixation in the A-questions is irrelevant for finding the answer. In this case, diagram fixation is therefore reflecting that participants are confused and fixate on a diagram actually comprising a distractor.

(3) Trial number shows a (non-linear) significant effect (p-value < 0.0001) on diagram-fixation time. Moreover, the diagram-fixation model shows an interaction between the explanatory variables trial number and question type, with a significant difference (p-value 0.0418) between the conditions DA and A, and also a significant (non-linear) difference (p-value 0.0045) between the conditions D and A.

Compared to the response-time model, the significant difference in trial-number effect for DA-questions (compared to A-questions) reappears. In addition, a difference in trial-number effect for D-questions (compared to A-questions) emerges (see figure 7.5).

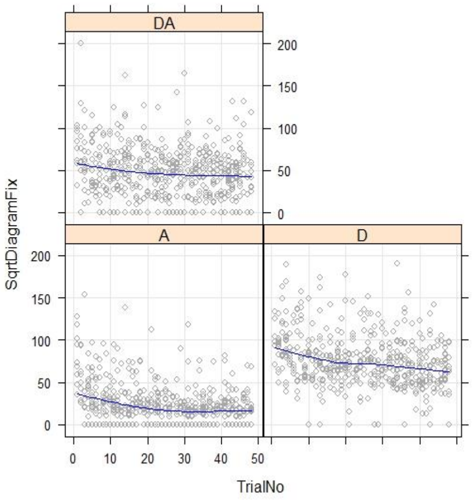


Figure 7.5: Interaction between question type and trial number on diagram-fixation time. Scatterplots for the raw data on the potential interaction between the three-level question type (DiagramArticle, Article and Diagram) and trial number (TrialNo) on the dependent variable diagram-fixation time.

The trial-number effect for A-questions differs significantly from the DA- and D-questions (see the different profile of the A-curve in the bottom left panel of figure 7.5), reflecting a different learning effect on the A-question. It appears that participants learn to ignore the diagram-distractor on the A-question and reduce their diagram-fixation time. As the experiment proceeds, participants reduce their diagram-fixation time on the D-question reflecting lower difficulty, while the evolution of diagram-fixation time on the DA-questions is very weak (almost horizontal) reflecting lower preference for diagram use.

As with the correctness and response-time models, the diagram-fixation model shows no expertise effects. Overall, the trial-number effect is significant but differing across question types, which we may interpret as a learning effect reflecting that participants increase performance (depth and speed on correct answers) as the experiment proceeds.

7.5.2 Discussion of diagram fixation

In my view, entry modes causing longer fixations are due to increased processing and difficulty for that entry mode. This is likely to happen only if that entry mode is indispensable for finding the answer (see table 7.5), but the increased diagram fixation could also reflect increased (effortless) diagram use, if the answer can be found in both entry modes.

In the same way, as low response time reflects speed, I expected low diagram-fixation time to reflect a deep problem representation (depth) among participants, who have a large domain-specific knowledge and therefore are capable of retrieving answers quickly. However, as we saw in the correctness and response-time models (see the discussion of reduced, absent and reversed expertise effects in section 7.3.2), none of the expertise variables, including the self-rated domain-specific expertise, are significant predictors. If the usability of the dual-entry mode is a question of outcome and diagram fixation time, no user adaption regarding expertise is necessary, rather users adapt to the diagram entry (learning effect).

7.6 Conclusion

The dual-entry experiment produced a very high percentage of correct response, hence I conclude that overall domain-specific knowledge in the taxation domain can be transferred to target users by means of a complementary dual-entry mode.

In the response-time model, the D-questions require longer response times compared to the A- and DA-questions, however, the experiment shows learning effects (trial-number effects) for all three question types. In the diagram-fixation model, participants show learning effects, which interact with question type, in particular, diagram-fixation time is reduced on the D-questions as the experiment proceeds. I conclude that high correctness and significant learning effects reflect that target users adapt to the dual-entry mode, which means that diagrams should be part of the interface in combination with the article.

I conclude that the performance models reflecting correctness, speed in terms of response time and depth in terms of diagram fixation in relevant questions show no expertise effects. This opens for three interpretations:

Reduced expertise effects from dual-entry modes: The dual-entry mode stimulus-pairs challenge, and potentially overload, the limited human cognitive capacity with redundant information. Absent expertise effects from biased expertise measures: It could be that the non-significant expertise effects are a matter the insufficient direct measures (chapter 5), as well as the indirect measures (chapter 6) of domain-specific expertise, which do not fully capture the complex dimensions and characteristics of expertise. It may also be a case of learning effects overriding potential expertise effects. Reversed expertise effects from the inflexibility of experts: experts apply (stereotypical) problem-solving strategies to uncommon tasks, while the non-experts may succeed in retrieving correct answers accidentally from the multiple-choice question format.

The result suggests that the non-adapted dual-entry mode prototype in the experiment is allowing users across all levels of expertise to adapt to the tool and acquire knowledge. From this benchmark further user adaption may be introduced, retested and evaluated in an iterative process. It is not possible that the experiments show facilitatory effects on non-expert performance, neutralized by inhibitory effects on expert performance, as no interactions involving expertise are significant.

Regarding the research theme, we may conclude that for the professional target users, the design of the low-fidelity dual-entry mode is usable (or will be after a few trials) i.e. users acquire knowledge without excessive response times, and the proposed design including the terminological ontology is a good starting point for an iterative design process.

Chapter 8: Future research

In this chapter, I discuss implications of my dissertation research (section 8.1). Moreover, I propose future research directions addressing the weaknesses of the chosen methods in this dissertation (section 8.2). Moreover, I discuss my research in relation to outstanding work addressed by the most important research areas: terminology and knowledge engineering (section 8.3) and usability engineering (section 8.4).

8.1 Results and implications

My dissertation research shows that domain-specific knowledge can be disseminated by means of the dual-entry modes, which comprise both concept articles and terminological ontologies, to all professional target users. Moreover, the research shows that performance is improved as the experiment proceeds, suggesting learning effects, and that target users adapt to the novel dual-entry modes, and we may conclude that terminological ontologies should be part of the user interface. However, the performance models showed no expertise effects.

The key results of the dissertation are the measures aimed at assessing the complex expertise variable and the use of terminological ontologies directly in term-bank interfaces. As the title suggests, I focus my research on knowledge dissemination based on terminological ontologies in term banks. However, the research also applies to the use of terminological ontologies in knowledge dissemination outside term banks. Below I discuss three key implications:

(1) Knowledge dissemination and cross-cultural barriers

As the dissertation title suggests, I base the knowledge dissemination on terminological ontologies visualized by dual-entry modes. Terminology is obviously not confined to terminology work, i.e. the representation in terminological ontologies. As discussed in chapter 2, terminology furthers the specialized discourse by ensuring unambiguous knowledge dissemination, in a variety of contexts crossing barriers of intra-domain cultures, e.g. the expert-to-non-expert dissemination between tax authorities and taxpayers, extra-domain cultures, e.g. expert-to-expert dissemination inside tax authorities or tax advisers, or supra-domain cultures, e.g. domain-specific dissemination from Danish to international English used in the EU (regional English) or OECD (global English).

(2) Term banks and discourse

My research is aimed at furthering the user-interface design of a term bank. However, the results also apply to other information tools than term banks. For instance, specialized dictionaries presenting domain-specific terminology by means of definitions and context examples may also benefit from the results on user-adaption demonstrated in this dissertation, despite the term-orientation. In addition, the insight into the expertise and expert performance of target users may be applied in revising other types of specialized communication constituting the discourse, to target readers.

(3) Terminological ontologies and system development

In my research, I focus on knowledge dissemination by means of terminological ontologies. However, terminological ontologies are not only used to represent knowledge and structure the entries of term banks. The formal conceptual modelling by means of formal feature specifications allows the content to be machine-readable and applicable for system development, e.g. as meta-data taxonomies or information storage (see e.g. Madsen, 2006).

8.2 Limitations

The key results are highly dependent on the chosen methods. It goes without saying that there are infinitely many ways to conduct research, but I have chosen the most relevant solutions given the practical and theoretical limitations. Below I discuss four points of improvements to the applied research strategy:

(1) Taxation domain

First, the taxation domain is not only about the revenue-collection discourse. A huge body of the taxation discourse is about tax compliance, and the international exchange of practices to ensure compliance has the potential of being supported by term banks and terminological ontologies. Second, terminological ontologies may contain a variety of relations, not only type relations. In the taxation domain, important aspects may benefit from part-whole relations e.g. to illustrate taxation structures of the revenue, or temporal relations e.g. to illustrate the many time constraints of the taxation domain. Third, it would be interesting to use different source language (SL) than Danish legal language, and compare that to other taxation systems, i.e. the alignment of taxation ontologies.

(2) Direct expertise measures

As concluded in chapter 5, the proposed expertise variables (participation, motivation, education and exposure) need to be improved to better reflect domain-specific expertise. Moreover, the direct self-rating measures of domain-specific expertise should be applied in all three dimensions of the expertise.

(3) Indirect expertise measures

As concluded in chapter 6, the proposed representative tasks and the evaluation of the performance of participants should be further developed to meet the deficiencies outlined by table 4.1. First, in the recalling task, the required number of terms to list should be extended, and we need to implement a time constraint. Second, in the categorizing task, the time should be recorded, and the definitions must be evaluated. Third, in the reading task follow-up questions could reflect the correct understanding of the text. Fourth, in the structuring task a time constraint must be implemented and a discussion with each participant about the mapping strategy may enrich the results.

(4) Dual-entry modes

First of all, the research results rest on the context and design of the dual-entry modes. It is quite possible that experiments cannot fully match the actual context of an information system (ecological validity), because that would require scenarios outlining the multiple-factor situation (environmental, organizational or technical). The location in an eye-tracking laboratory at a business school and in front of a remote eye tracker, where participants were asked to avoid head movements hardly matches the actual environmental circumstances of use of a term bank. Moreover, the scenario is merely on-screen and detached from an actual work situation of a newspaper journalist.

Second, we may refine the analysis by considering smaller AOIs, which would allow us to evaluate, whether participants are looking in the right area of each entry mode. In the concept articles, we may further investigate if participants are fixating in the relevant information category. For instance, we may ask ourselves, where the primary fixation of the diagram is taking place: in the top, or at the bottom.

Third, we may include scan-path data in the analysis to reflect the chosen search routes of participants (see the bottom panels of figure 4.4). Long response times are not necessarily a case of long fixation time. Reponse time may also be a matter of browsing the dual-entry mode intensively resulting in long scan paths. In the latter case, the participant is not able to locate the answer in the dual-entry mode, while in the former case, we use the eye-to-mind hypothesis to conclude that the participant is not able understand the content, he is fixating on. It should be noted that the experimental design, where multiple-choice questions are placed above the stimulus-space (see figure 4.3), will necessarily produce long scan paths as participants need to check a retrieved answer with the available answers prior to the keying-in of the answer, which terminates the response time.

8.3 Terminology and knowledge engineering issues

When choosing research themes and pursuing a research strategy to answer the questions, I faced many possibilities as well as constraints, which had to be disregarded, but should be included in future analysis to complete the picture of user-adapting term banks. In the area of terminology and knowledge engineering, two of the most important directions are discussed below:

(1) Multimedia term banks

The terminological ontologies constitute graphs, which may be considered as a type of multimedia (image), but the full advantage of the multimedia potential is outstanding.

User-oriented lexicography and terminology should not only consider the adaption of the user-interface, users should also benefit from multimedia content. Rylova (2011) emphasizes the one-database approach to the compilation dictionaries published and adapted into multiple formats, where crucial questions become optimal size, user-friendly layout, user-interaction across formats. De Shryver (2003, p. 188) subsumes the dictionary of the third millenium by stating that “the direction in which electronic lexicography is moving is exactly this: towards more content, more flexibility and customisation, more user-friendliness, better access and more connectivity with other sources of knowledge.”

De Schryver (2003, p. 146) accounts for the lexicographer's dreams including the user access to electronic dictionaries (EDs) and underlines that “without truly implementing fully integrated hypermedia access structures, EDs aren't really very different from their paper counterparts.” We may conclude that the same thing applied for term banks, i.e. multimedia should be included.

Koplenig (2011) asked users about crucial aspects of usability and concluded that users rate the distinctive characteristics of EDs (i.e. multimedia and user-adaption) partly unimportant, which conflicts with the request for multimedia elements and user-adaptive interfaces. In particular, the users appear old-fashioned rating classic criteria like reliability and clarity highest. However, future research on term banks will show whether terms banks benefit from the inclusion of multimedia content.

(2) Target-user validation

In building term banks manually, the iterative process of terminology work, includes the merging of doublets (several term entries cover the same concept and should be merged). Current terminology research is focusing on automatic validation (control for consistency). However, domain experts are still needed to complete the validation (DanTermBank, 2015c).

It should be noted that semantic inconsistency is not necessarily undesirable, if we ask the users. In the Swedish national term bank (Rikstermbanken), a survey gives the quite surprising result that users prefer to investigate the doublets and would not welcome any merge of entries performed by the term bank developers.

Perhaps this reflects target users' need to access information as a sound basis for establishing the terminology fitting their particular context.

The subjective views revealed by participants' concept maps demonstrated differing views rather than inconsistent structuring. Future research may be enriched by asking participants to validate and perhaps correct the overall terminological ontologies.

8.4 Usability engineering issues

Usability engineering is a methodical "engineering" approach to user interface design and evaluation involving practical, systematic approaches to developing requirements, analyzing a usability problem, developing proposed solutions, and testing those solutions (Usability First, 2015). In the area of usability engineering, three of the most important directions for future work are discussed below:

(1) User-adaption

First, target users were limited to professionals. The dissertation research generalizes from the sample of professionals to the population of professionals. However, the target users of a national term bank are also non-professionals (i.e. children, students and retired people). Non-professionals are also important target users, because the retired people constitute a considerable group, who may need a term bank to understand the information from public authorities and the increased demands for digital self-service. The students may use the term bank to prevent domain loss.

(2) Prototyping

I conducted my research without a running prototype of the term bank, which should be included in future work. In particular, I should design user-adapted interfaces based on the results of this dissertation, and then test the redesign on participants with relevant level of expertise.

The term-bank prototype should allow users to full access and search in the tool. Term banks may benefit from the research on the search and interaction features in dictionaries. De Schryver (2003, p. 173) suggests that “[k]eying in, copy-and-pasting, and mouse clicking (…one only needs to position the mouse pointer over (and in some versions also single-click) words onscreen, after which the relevant dictionary article or articles pop up in a window) are today's most frequent actions used to uncover the data in an ED.” Moreover De Schryver (2003, p. 178) states that “[t]o speed up the look-up process even more, an increasing number of EDs use “focus-in typing”, aka “incremental search”, whereby a list of candidates is shown (and narrowed down) as one is typing. From the moment one sees the item one is looking for, one can simply select it […]”.

In the future prototype of the term bank, it should be possible to unfold the images further. Lew (2010) investigate three types of dictionaries from plain-entry (without navigational assistance) to the inclusion of entry-menus showing significantly faster performance. Therefore, the concept articles should contain further information inside the categories shown, but they should also be expandable with additional categories. Concept diagrams should allow the user to navigate across the terminological ontology, horizontally and vertically, as well as expanding the number of concepts and relations displayed (zooming out). These potential navigational aids were not part of the static dual-entry modes, but I chose not to visually mark these features (for instance by indicating with a “+” that the visualization was expandable), which in some cases led to some confusion, unfortunately. In particular, the graphical entry-mode displaying the conceptual diagram was prone to confusion, perhaps the best example illustrating this problem is the verbal response by one participant (no. 9): “Surely, more duties exist than shown in the diagram?!” (*Der er da flere afgifter, end diagrammet viser!*). However, the aim was to keep entry modes as simple and to the point as possible without too many features, and as the pilot studies did not reveal much confusion changes were deemed unnecessary.

(3) Usability methods

Usability metrics constitute “formal measurements that are used as guides to the level of usability of a product. Metrics include how fast a user can perform a task, number of errors made on a task, learning time, and subjective ratings.” (Usability First, 2015). The performance models predicting user behaviour in the dual-entry mode experiments were guided by expertise characteristics, i.e. correctness, speed and depth. Future research may benefit from additional metrics, especially the learning time and subjective ratings, as well as the assessment of associated levels to indicate satisfactory usability.

Think-aloud is “a technique in user testing where users are asked to speak their thoughts as they perform a task, while the focus in user testing is primarily on how effectively a user performs the required tasks (and not on how users believe they are performing), verbalizations are quite useful in understanding mistakes that are made and getting ideas for what the causes might be and how the interface could be,” (Usability First, 2015).

Asking participants to think-aloud, while doing the dual-entry mode experiment in front of the eye tracker would have produced valuable insight into users immediate perception, which is not necessarily reflected in the retrospective interview, simply because participants have forgotten their reaction to each of the 48 questions. However, think-aloud pose a serious risk of overloading participants. It should be noted that some participants burst their views out, which was collected as additional data, however, most participants keep very quiet and focused.

Despite the de-contextualization, the key issue is to ensure as natural user behaviour as possible avoiding unnecessary confusion, which generalizes to settings outside the experiment. This is sometimes referred to as external validity (Usability First, 2015). At one future point in time, the term bank will be implemented, and then it will be relevant to conduct surveys reflecting performance and satisfaction of target users. This type of research allows the participants to stay in their natural environment using their own computer without any eye trackers, which increases the ecological validity of the research. In the DanTermBank project so-called user-scenarios were a type of survey conducted at primary and secondary schools, where a group of school children had access to a term bank prototype and a control group were left with only internet access (DanTermBank, 2015d).

8.5 Conclusion

I conclude that my research applies to knowledge dissemination in other contexts, to information tools different from term banks, and to other applications of terminological ontologies.

Moreover, I conclude that adjustments to the chosen domain and methods of this dissertation research should alleviate its limitations, i.e. the chosen taxation discourse should be expanded, the direct measures should reflect the complex expertise, the indirect measures of expertise should reflect all expertise characteristics, and the eye-tracking analysis should be more detailed.

Finally, I conclude that the chosen research methods are based on the most relevant research fields, i.e. terminology and knowledge engineering and usability engineering. Regarding the former, outstanding research is the inclusion of multimedia in the user-interface and target-user validation of terminology, as well as user-adaption to different target users, by means of different user characteristics other than expertise, by using a running a user-adapted prototype with dynamic functionalities and by applying more usability methods e.g. think-aloud and user surveys.

To sum up, future work should examine the optimization of usability on the mentioned search, visualization and navigation features, without disregarding the fundamental principles of terminology.

Appendix

Appendix I: Systematic list of concepts

Table I.1: Example of the concept clarification process for "personal income tax" (*personskat*). 1: The notation following the structure of the concept system in figures II.1-II.19. 2: The concept is the content of the term list extracted from the source. The characteristics are stated by "(attribute: value)". 3: Terms are stated in Danish for the top-three sources (Ministry of Taxation (SKM), Ministry of Finance (FFL), Statistics Denmark (DST) and in English for the bottom-three sources (Statistics Denmark (DST), Eurostat and OECD) as the publication from the former is bilingual. 4: The page numbers in the sources refer to where the revenue figures are found. 5: Acts are numbered. Some are historic. 6: Revenues are for the year 2010 and measured in billion Danish Kroner. Revenues in italics are the taxes to the central government.

Systematic list (1)	**Conceptual content (2)**	**Term extraction of lexical units (3)**	**Source set (4)**	**Act (legal information) (5)**	**Revenue (financial information) (6)**
1	skat				
1.7	direkte skat (collection form: With the taxpayer)	lov om gensidig administrativ bistand i sager om direkte og indirekte skatter mellem stater, der er medlem af Europarådet eller OECD	SKM	LOV nr 132 af 26/02/1992	N.a
1.7.1	personskat (taxpayer: liable persons)	bekendtgørelse af lov om indkomstskat for personer m.v. personskatteloven	SKM	LBK nr 143 af 08/02/2011	n.a.
		personskatter	FFL (2012) (p. 296)	LB 143 2011	*227.5*
		personlige indkomstskatter personbeskatning personskatteloven lov om indkomstskat for personer mv. de almindelige personlige indkomstskatter af kommunerne personlige indkomstskat til staten, kommunerne og amtskommunerne (til og med 2006) indkomstskatterne omfatter skatter fra personer, selskaber o.l. den personlige indkomstbeskatning den personlige indkomstskat personrelaterede skatter personskatteområdet personskat	DST (2012) (p. 49)		427.1 *232.9*
		personal income tax personal taxation personal tax rate			
		personal income personal taxation	Eurostat (2012) (p. 76)		429.8 (57.3 €)

		personal income tax	OECD (2011) (p. 22)		n.a.

Table I.2: Systematic list of 94 concepts and their attributes constituting the Danish tax system. 1: The notation following the structure of the concept system. Concepts not part of the structure are marked 'n.a.' and written in grey. 2: The concept including the attribute-value pair stated in source language Danish. 3: The concept including the attribute-value pair stated glossed in English. 4: The concepts chosen for the experiments (see appendix D) are marked. The concepts in the experimentation have been through a concept clarifying process (see chapter 2), whereas the concepts *not* part of the experimentation are merely glossed and in italics: If the term can be found in Gyldendal dictionaries, the term is stated with the reference in parenthesis. If the term cannot be found in Gyldendal dictionaries, a 'n.a.' is in parentheses. *: Examples of term formation.

Systematic list (1)	**Concept including attribute-value pair (2)**	**Glossary (3)**	**Part of experimentation (4)**
1	skat	’tax’	Yes PUBLIC FINANCES: income
1.1	statsskat (modtager af provenu: staten (national))	’central government tax’ (receiver of revenue: state (national))	Yes RECEIVER: central government
1.2	kommunal skat / kommuneskat (modtager af provenu: kommunen (lokal))	’local state tax’ (receiver of revenue: municipality (local))	Yes RECEIVER: local government
1.3	sundhedsbidrag (modtager af provenu: regioner (regional))	’healthcare contribution’ (receiver of revenue: regions (regional))	Yes RECEIVER: local authorities
1.4	bidrag til sociale ordninger (modtager af provenu: sociale kasser og fonde)	*contribution to social schemes (n.a.)* (receiver of revenue: social foundations)	No
1.5	kirkeskat (modtager af provenu: provstiet (folkekirken))	*church tax (Gyldendal)* (receiver of revenue: deanery (folkekirken))	No
1.6	EU-skatter (modtager af provenu: EU (over-national))	*EU taxes (n.a.)* (receiver of revenue: EU (supranational))	No
1.7	direkte skat (opkrævningsform: personligt hos skatteyderen)	’direct tax’ (collection form: directly with the taxpayer)	Yes COLLECTION FORM: direct with the taxpayer
1.7.1	personskat (skatteyder: skattepligtige personer)	’personal income tax’ (taxpayer: liable persons)	Yes TAXPAYER: taxable persons

1.7.1.1	forskudsskat (selvstændig erhvervsdrivende: Nej (lønmodtager))	’provisional tax’ (self-employed: no (wage earner))	Yes PAYMENT: continual, in advance of final settlement.
1.7.1.2+ 1.7.2.1	virksomhedsskat (selvstændig erhvervsdrivende: Ja)	’corporation tax’ (self-employed: yes)	Yes TAXPAYER: self-employed persons.
1.7.2	selskabsskat (skatteyder: skattepligtige selskaber)	’corporate tax’ (taxpayer: liable companies)	Yes TAXPAYER: taxable companies
1.7.2.2	acontoskat (selvstændig erhvervsdrivende: Nej (selskab))	’prepaid tax’ (self-employed: no (company))	Yes PAYMENT: continuous, in advance of final settlement
1.7.3	fonds- og foreningsskat (skatteyder: skattepligtige fonde og foreninger)	*tax on funds and associations (n.a.)* (taxpayer: taxable funds and associations)	No
1.7.4	kulbrinteskat (skatteyder: indvindingsvirksomheder (til staten))	*carbon tax (Gyldendal)* (taxpayer: extraction corporations (to central government))	No
1.7.5	statslig indkomstskat (grundlag for direkte skat: skattepligtig indkomst (til staten))	’state income tax’ (direct tax base: taxable income (to central government))	Yes RECEIVER: central government
1.7.5.1	topskat (skattesats: sats på den del af den skattepligtige indkomst, som overstiger en beløbsgrænse (progressiv skala))	’top-bracket tax’ (tax rate: rate applicable only on the part of the taxable income exceeding a certain limit (progressive scale))	Yes INCOME LIMIT: high
n.a.	mellemskat	’middle-bracket tax’	Yes INCOME LIMIT: medium
1.7.5.2	bundskat (skattesats: sats på hele den skattepligtige indkomst (proportional skala))	’lower-bracket tax’ (tax rate: rate applicable on the entire taxable income (proportional scale))	Yes INCOME LIMIT: low
1.7.6	kommunal indkomstskat / kommuneskat (grundlag for direkte skat: skattepligtig indkomst (til kommunen))	’local income tax’ (direct tax base: taxable income (to local government))	Yes RECEIVER: local government
1.7.7	arbejdsmarkedsbidrag (grundlag for direkte skat: indkomst indberettet af skatteyders arbejdsgiver(e) (til staten))	*labour-market contribution (Gyldendal)* (direct tax base: income reported by the employer (to central government))	No
1.7.8	kapitalindkomstskat (grundlag for direkte skat: kapitalindkomst, netto)	*capital gains tax (Gyldendal)* (direct tax base: net capital gains)	No
1.7.8.1	skat af CFC-indkomst (indkomsttype: CFC-indkomst)	*CFC-income tax (n.a.)* (income type: CFC-income)	No

1.7.9	aktieindkomstskat (grundlag for direkte skat: aktieindkomst (til staten))	*equity income tax (n.a.)* (direct tax base: equity income (to central government))	No
1.7.9.1	aktieudbytteskat (indkomsttype: aktieudbytte)	*dividend tax (n.a.)* (income type: dividend)	No
1.7.9.2	aktieavanceskat (indkomsttype: avance fra aktiehandel)	*share premium tax (Gyldendal)* (income type: share premium)	No
1.7.10	ejendomsværdiskat (grundlag for direkte skat: værdi af ejendom (til staten))	’property value tax’ (direct tax base: property value (to central government))	Yes TAX BASE: property value assessment
1.7.11	ejendomsskat (grundlag for direkte skat: værdi af grund (til kommunen))	’land tax’ (direct tax base: land value (to local government))	Yes TAX BASE: land value
1.7.11.1	grundskyld (obligatorisk: ja)	’land value tax’ (obligatory: yes)	Yes OBLIGATORY: yes
1.7.11.2	dækningsafgift (obligatorisk: nej (offentlige ejendomme))	’reimbursement duty’ (obligatory: no (public sector properties)	Yes OBLIGATORY: no
1.7.12	afgift af dødsboer og gaver (grundlag for direkte skat: indkomst fra boskifte, arv og gaver)	*tax on estate of deceased persons or gifts (n.a.)* (direct tax base: income from estate of deceased persons, inheritance and gifts)	No
1.7.13	frigørelsesafgift (grundlag for direkte skat: avance fra zoneændring på ejendom)	*property release tax (Gyldendal)* (direct tax base: profit from changing zone of property)	No
1.7.14	pensionsafkastskat (grundlag for direkte skat: afkast på pension)	*pension savings return tax (Gyldendal)* (direct tax base: return on pension savings)	No
1.7.15	udligningsskat (grundlag for direkte skat: pensionsudbetalinger (til staten))	*Term: equalization tax (n.a.)* *Definition: local tax or rate paid in one area of local government as a contribution to another (Gyldendal)* (direct tax base: pension payout (to central government))	No
1.8	indirekte skat (opkrævningsform: på transaktionen)	’indirect tax’ (collection form: on the transaction)	Yes COLLECTION FORM: indirectly on the transaction
1.8.1	punktafgift (grundlag for indirekte skat: forbrug af varer og tjenester (punkter))	’excise duty’ (indirect tax base: consumption of goods and services (excises))	Yes TAX BASE: consumption of goods and services

n.a.	energiskat	’energy tax’	Yes PURPOSE: limiting environmentally damaging energy consumption
n.a.	grøn afgift	’green tax’	Yes PURPOSE: limiting environmentally damaging use and consumption
1.8.1.1	afgift af motorkøretøj (formål: begrænse forbrug af miljøskadelige motorkøretøjer)	’motor vehicles tax’ (purpose: limit use of damaging vehicles)	Yes PURPOSE: limiting environmentally damaging use of vehicles
1.8.1.1.1	grøn ejerafgift (grundlag for indirekte skat: motorkøretøjets brændstofforbrug (km pr. liter))	’green vehicle excise duty’ (indirect tax base: fuel consumption of the vehicle (km per litre))	Yes TAX BASE: fuel consumption.
1.8.1.1.2	vægtafgift (grundlag for indirekte skat: motorkøretøjets art og egenvægt)	’weight tax’ (indirect tax base: type and weight of the vehicle)	Yes TAX BASE: the type and weight of the vehicle
1.8.1.1.3	registreringsafgift (grundlag for indirekte skat: værdi af motorkøretøj)	’motor vehicle registration duty’ (indirect tax base: value of the vehicle)	Yes TAX BASE: the value of the vehicle
1.8.1.1.4	afgift af vejbenyttelse (grundlag for indirekte skat: lastvognens udstødningsklasse og størrelse)	’road charges’ (indirect tax base: truck size and exhaust-emission class)	Yes TAX BASE: truck size and exhaust-emission class
1.8.1.1.5	afgift af ansvarsforsikring (grundlag for indirekte skat: præmien på motorkøretøjets ansvarsforsikring)	’duty on third-party liability insurance’ (indirect tax base: premium on third party liability insurance of the motor vehicle)	Yes TAX BASE: premium on third party liability insurance of the motor vehicle
1.8.1.1.6	udligningsafgift (egenskab: dieseldreven)	*countervailing charge (Gyldendal)* (property: diesel-powered)	No
1.8.1.1.7	partikeludledningsafgift (egenskab: dieseldrevne uden partikelfilter)	*particulate discharge duty (n.a.)* (property: diesel-powered without particle filter)	No
1.8.1.1.8	privatbenyttelsesafgift (egenskab: gulpladebil, som benyttes privat)	*duty on privately used vehicles (n.a.)* (property: vehicle with yellow licence plates, used privately)	No

1.8.1.2	energiafgift (formål: begrænse miljøskadeligt energiforbrug)	’energy duty’ (purpose: limit damaging energy consumption)	Yes PRODUCT: environmentally damaging energy or fossile fuels Or TAX BASE: environmentally damaging energy consumption Or PURPOSE: limiting environmentally damaging energy consumption
1.8.1.2.1	afgift af kul (energikilde: kul)	*charge levied on coal (Gyldendal)* (energy source: coal)	No
1.8.1.2.2	afgift af elektricitet (energikilde: elektricitet)	*charge levied on electricity (Gyldendal)* (energy source: electricity)	No
1.8.1.2.3	afgift af naturgas (energikilde: naturgas, bygas)	*duty on natural gas (n.a.)* (energy source: natural gas, town gas)	No
1.8.1.2.4	afgift af visse olieprodukter (energikilde: olieprodukter, autogas, flaskegas)	*charge levied on certain oil products (n.a.)* (energy source: oil products, auto gas, bottled gas)	No
1.8.1.8.5	afgift af benzin (energikilde: benzin)	*duty on petrol (Gyldendal)* (energy source: petrol)	No
1.8.1.3	miljøafgift (formål: begrænse miljøbelastning)	’environmental tax’ (purpose: limit environmental damage)	Yes TAX BASE: environmentally damaging substances or scarce natural resources Or PURPOSE: limiting environmentally damaging products and contents
1.8.1.3.1	kuldioxidafgift (miljøbelastende indhold: kuldioxid CO2)	’carbon dioxide tax’ (damaging content: carbon dioxide CO2)	Yes CONTENT: carbon dioxide
1.8.1.3.2	afgift af svovl (miljøbelastende indhold: svovl og svovldioxid SO2)	’duty on sulphur’ (damaging content: sulphur and sulphur dioxide SO2)	Yes CONTENT: sulphur
1.8.1.3.3	afgift af kvælstofoxider (miljøbelastende indhold: NO2-ækvivalenter)	’duty on nitrogen oxides’ (damaging content: NO2-equivalents)	Yes CONTENT: nitrogen oxides

1.8.1.3.4	afgift af kvælstof (miljøbelastende indhold: kvælstof i ammoniak, kalium- og cadmiumnitrat, amminiumchlorid og gødningsstoffer)	*duty on nitrogen (n.a.)* (damaging content: nitrogen in ammonia, potassium nitrate, amminium chloride and fertilizer)	No
1.8.1.3.5	CFC-afgift (miljøbelastende indhold: CFC)	*duty on CFC (n.a.)* (damaging content: CFC)	No
1.8.1.3.6	afgift af klorede opløsningsmidler (miljøbelastende indhold: klorerede opløsningsmidler)	*duty on chlorinated solvents (n.a.)* (damaging content: chlorinated solvents)	No
1.8.1.3.7	afgift af bekæmpelsesmidler (miljøbelastende indhold: bekæmpelsesmidler)	*duty on pesticide (n.a.)* (damaging content: pesticides)	No
1.8.1.3.8	afgift af visse vækstfremmere (miljøbelastende indhold: antibiotika, vækstfremmende stoffer, der anvendes som tilsætningsstoffer til foderstoffer)	*duty on specific growth stimulants (n.a.)* (damaging content: antibiotics, growth promoting substances, used as additives in feedstuffs)	No
1.8.1.3.9	afgift af mineralsk fosfor (miljøbelastende indhold: mineralsk fosfor i foderfosfat)	*duty on mineral phoshorus (n.a.)* (damaging content: mineral phosphorus in feed phosphate)	No
1.8.1.3.10	afgift af pvc og phathalater (miljøbelastende indhold: pvc)	*duty on PVC and phathalates (n.a.)* (damaging content: PVC)	No
1.8.1.3.11	emballageafgift (miljøbelastende produkt: salgsemballage)	*packaging tax (Gyldendal)* (damaging product: sales packaging)	No
1.8.1.3.12	afgift af affald (miljøbelastende produkt: affald til deponering)	*waste fee (Gyldendal)* (damaging product: waste)	No
1.8.1.3.13	afgift af spildevand (miljøbelastende produkt: spildevand, der udledes eller nedsiver)	*effluent charge (Gyldendal)* (damaging product: effluent water, discharged or leaking)	No
1.8.1.3.14	afgift af nikkel/cadmium batterier (miljøbelastende produkt: nikkel/cadmium akkumulatorer)	*duty on nickel cadmium batteries (n.a.)* (damaging product: nickel cadmium accumulators)	No
1.8.1.3.15	afgift af glødelamper (miljøbelastende produkt: glødelamper, elektriske sikringer)	*duty on glow lamps (n.a.)* (damaging product: glow lamps, electric fuses)	No
1.8.1.3.16	afgift af råstoffer (knap naturressource: råstoffer)	*duty on raw materials (n.a.)* (scarce resource: raw materials)	No
1.8.1.3.17	afgift af ledningsført vand (knap naturressource: ledningsført vand)	*duty on piped water (n.a.)* (scarce resource: piped water)	No
1.8.1.4	afgift på spil (formål: begrænse forbrug af spil)	’duty on gaming’ (purpose: limiting demand for games)	Yes PURPOSE: limiting demand for games

1.8.1.4.1	afgift af gevinster ved lotterispil (spilletype: lotteri)	*duty on lottery win (n.a.)* (game type: lottery)	No
1.8.1.4.2	afgift af tipning (spilletype: tipning)	*duty on the pools (n.a.)* (game type: doing the pools)	No
1.8.1.4.3	afgift af spil ved væddeløb (spilletype: væddeløb)	*duty on track race (n.a.)* (game type: track race)	No
1.8.1.4.4	afgift af spilleautomater (spilletype: gevinstgivende spilleautomater)	*duty on gambling machines (n.a.)* (game type: gambling machines)	No
1.8.1.4.5	kasinoafgift (spilletype: spillekasino (bruttospilleindtægter))	*duty on casinos (n.a.)* (game type: gambling casinos (gross income))	No
1.8.1.5	afgift af nydelsesmidler* (formål: begrænse sundhedsskadeligt forbrug)	’duty on stimulants’ (purpose: limit unhealthy consumption)	Yes PURPOSE: limiting unhealthy consumption of food, drink and tobacco
1.8.1.5.1	afgift af spiritus (alkoholisk drikkevare: spiritus (ethanol))	*duty on spirits (Gyldendal)* (alcoholic beverage: spirits (ethanol))	No
1.8.1.5.2	afgift af vin (alkoholisk drikkevare: vin (ethanol))	*duty on wine (n.a.)* (alcoholic beverage: wine (ethanol))	No
1.8.1.5.3	afgift af øl (alkoholisk drikkevare: øl (alkohol))	*beer duty (Gyldendal)* (alcoholic beverage: beer (alcohol))	No
1.8.1.5.4	tillægsafgift af alkoholsodavand (alkoholisk drikkevare: alkoholsodavand)	*additional duty on alcopop (n.a.)* (alcoholic beverage: alcopop)	No
1.8.1.5.5	afgift af mineralvand (ikke-alkoholisk drikkevare: mineralvand, saft, most, læskedrikskoncentrat (liter))	*duty on mineral water (n.a.)* (non-alcoholic beverage: mineral water, juice, fruit juice, soft drink concentrate (liter))	No
1.8.1.5.6	afgift af kaffe (ikke-alkoholisk drikkevare: rå kaffe, brændt kaffe, kaffeekstrakt, kaffeerstatning (kilo))	*duty on coffee (n.a.)* (non-alcoholic beverage: raw coffee, roasted coffee, coffee extract, coffee substitute (kilo))	No
1.8.1.5.7	afgift af te (ikke-alkoholisk drikkevare: te, teekstrakt (kilo))	*duty on tea (n.a.)* (non-alcoholic beverage: tea, tea extract (kilo))	No
1.8.1.5.8	afgift af chokolade (vareområde: chokolade)	*duty on chocolate (n.a.)* (product area: chocolate)	No
1.8.1.5.9	afgift af konsumis (vareområde: is)	*duty on ice-cream (n.a.)* (product area: ice-cream)	No
1.8.1.5.10	afgift af mættet fedt (vareområde: fødevarer, som er primære kilder til mættet fedt)	*duty on saturated fat (n.a.)* (product area: primary food source for saturated fat)	No

1.8.1.5.11	afgift af andre tobaksvarer* (vareområde: kadusskrå, snus, anden røgfri tobak, cigaretpapir)	*duty on other tobacco (n.a.)* (product area: chewing tobacco, snuff, smokeless tobacco, cigarette paper)	No
1.8.1.5.12	tobaksafgift (vareområde: tobak)	*duty on tobacco (n.a.)* (product area: tobacco)	No
1.8.2	moms (grundlag for indirekte skat: merværdi (momspligtig virksomhed)	*value-added tax (Gyldendal)* (indirect tax base: added value (taxable business))	No
1.8.3	lønsumsafgift (grundlag for indirekte skat: lønsum (momsfri virksomhed))	*payroll tax (Gyldendal)* (indirect tax base: payroll (business exempted from value-added tax)	No
1.8.4	told (grundlag for indirekte skat: værdi af import eller eksport af varer og tjenester)	*customs duty (Gyldendal)* (indirect tax base: value of import or export of goods and services)	No
1.8.5	tinglysnings- og stempelafgift (grundlag for indirekte skat: værdi af tinglysning (offentlig registrering))	*registration and stamp duty (n.a.)* (indirect tax base: value of public registration)	No
1.8.5.1	stempelafgift på skadesforsikring (type: skade)	*stamp duty on general insurance (n.a.)* (insurance type: general insurace)	No
1.8.6	forsikringsafgift (grundlag: forsikingssum)	*duty on insurance (n.a.)* (tax base: insured sum)	No
1.8.6.2	afgift af lystfartøjsforsikring (type: lystfartøjer)	*duty on pleasure-craft insurance (n.a.)* (insurance type: pleasure craft)	No

Appendix II: Terminological ontology

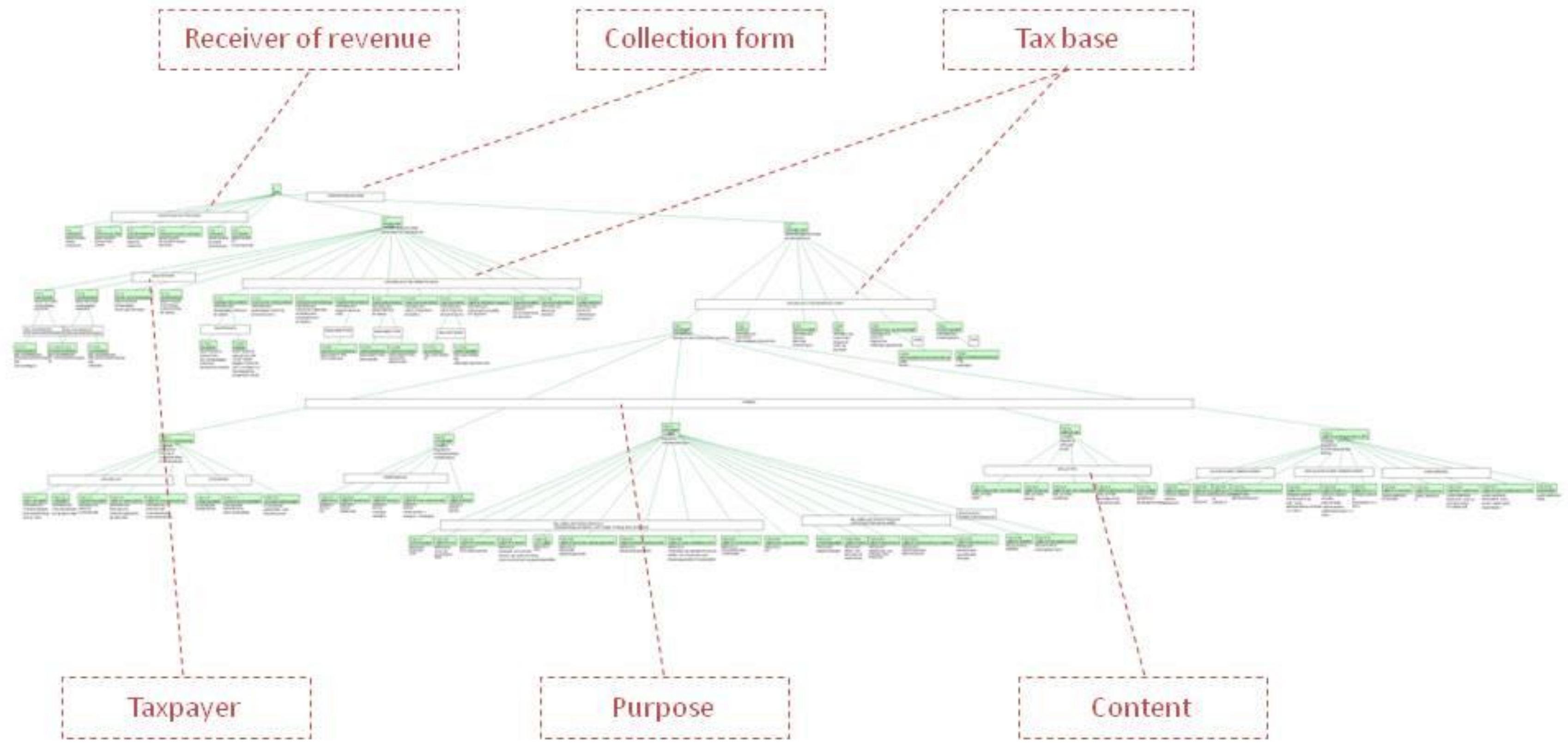


Figure II.1: Diagram visualizing the terminological ontology for the Danish tax system comprising the concepts identified in table I.2. The primary subdivision criteria are shown in boxes. For closer views of the system components see figures II.2-II.19.

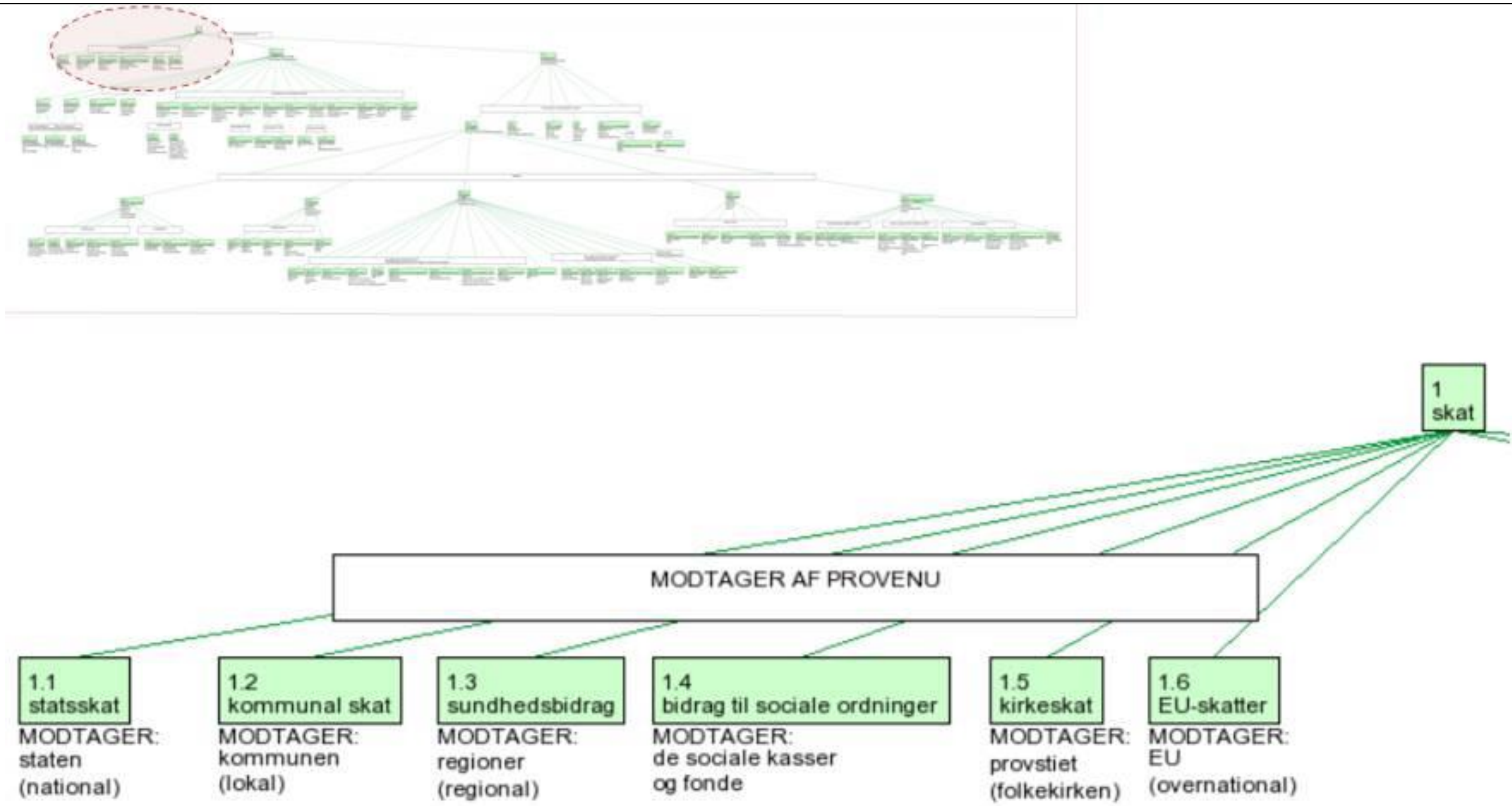


Figure II.2: Terminological ontology extract concerning the concept *skat* (**1**) and the subordinates **1.1-1.6** under the subdivision criteria: *MODTAGER AF PROVENU*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

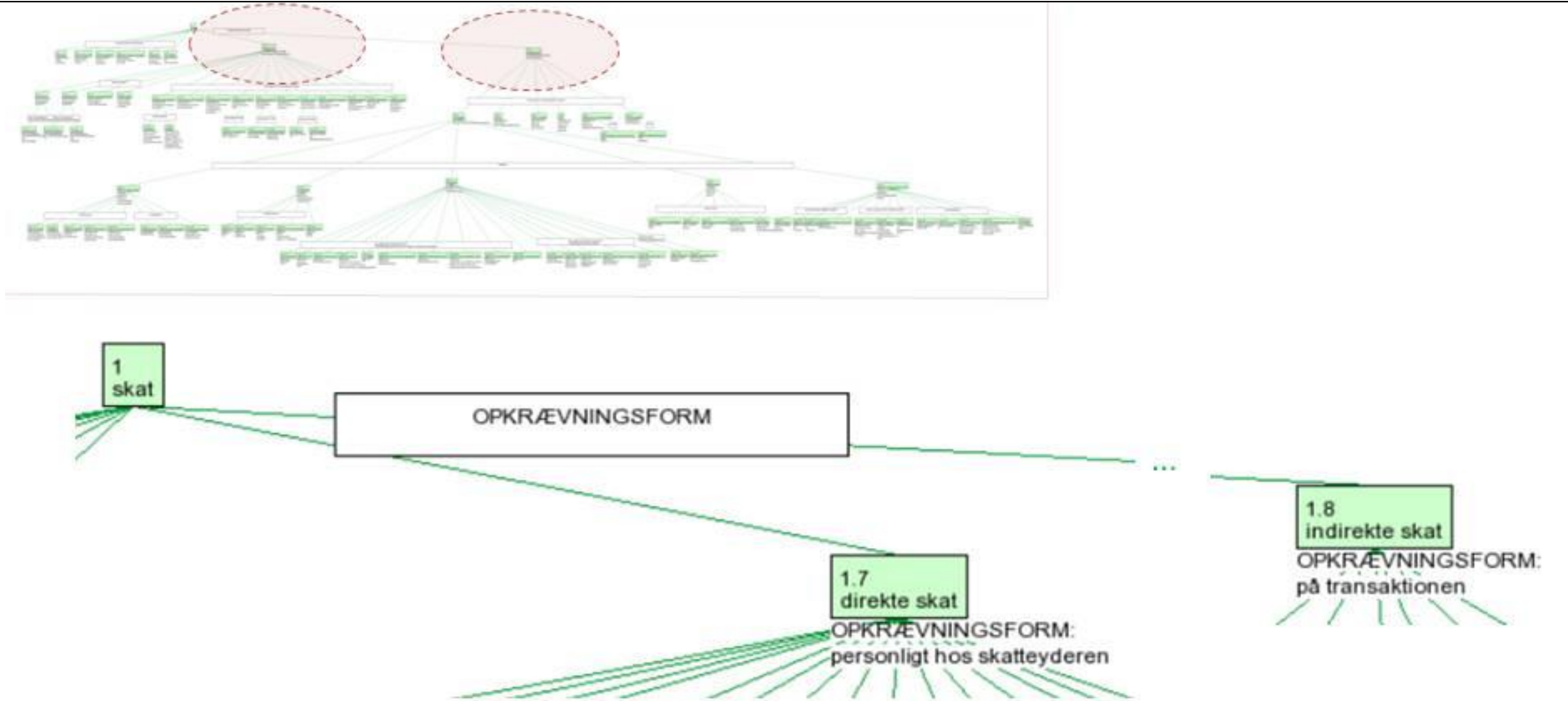


Figure II.3: Terminological ontology extract concerning the concept *skat* (**1**) and the subordinates **1.7-1.8** under the subdivision criteria: *OPKRÆVNINGSFORM*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

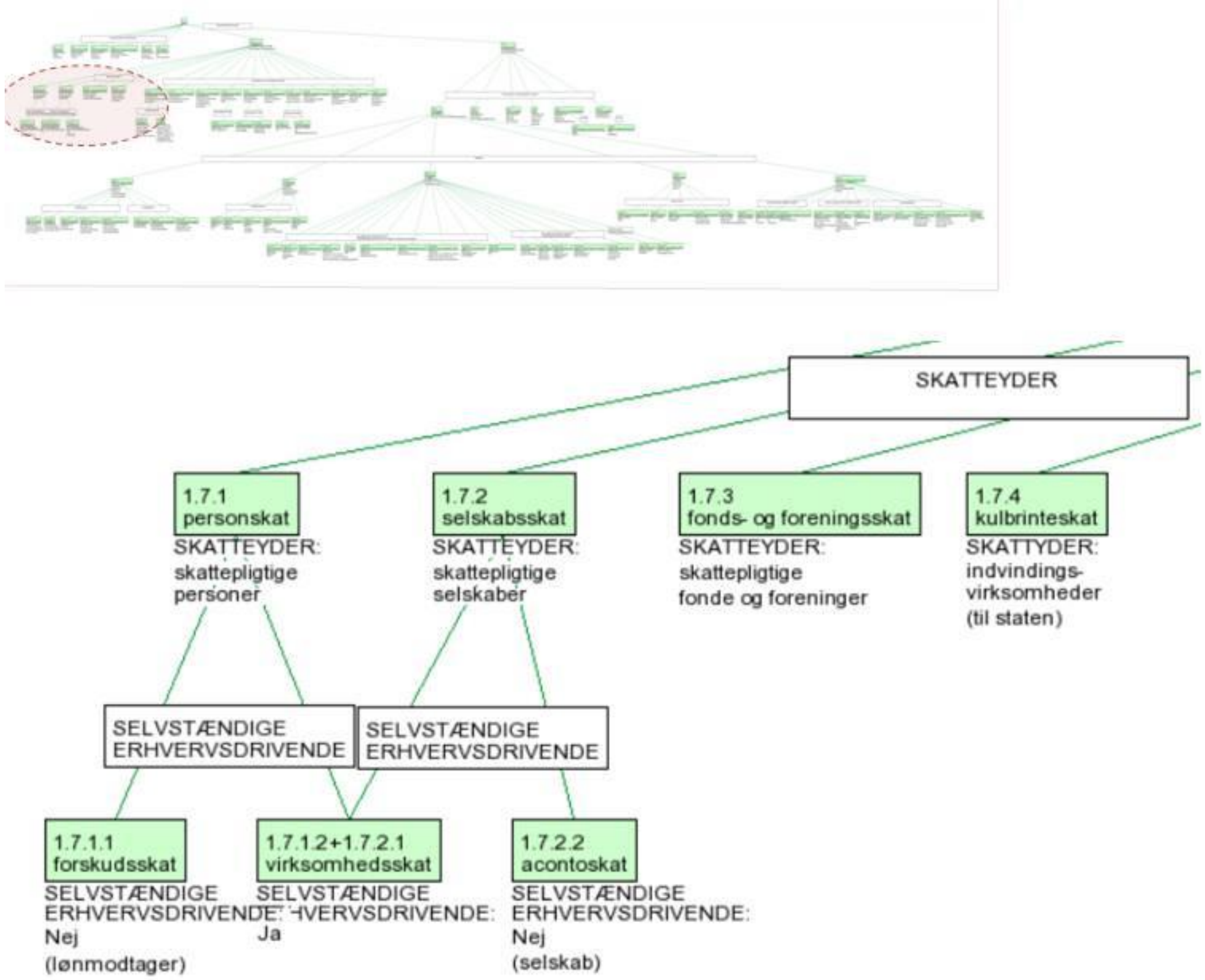


Figure II.4: Terminological ontology extract concerning the concept *direkte skat* (**1.7**), the subordinates **1.7.1-1.7.4** under the subdivision criteria: *SKATTEYDER*, and the sub-subordinates **1.7.1.1-1.7.2.2** under the subdivision criteria: *SELVSTÆNDIG ERHVERVSDRIVENDE*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

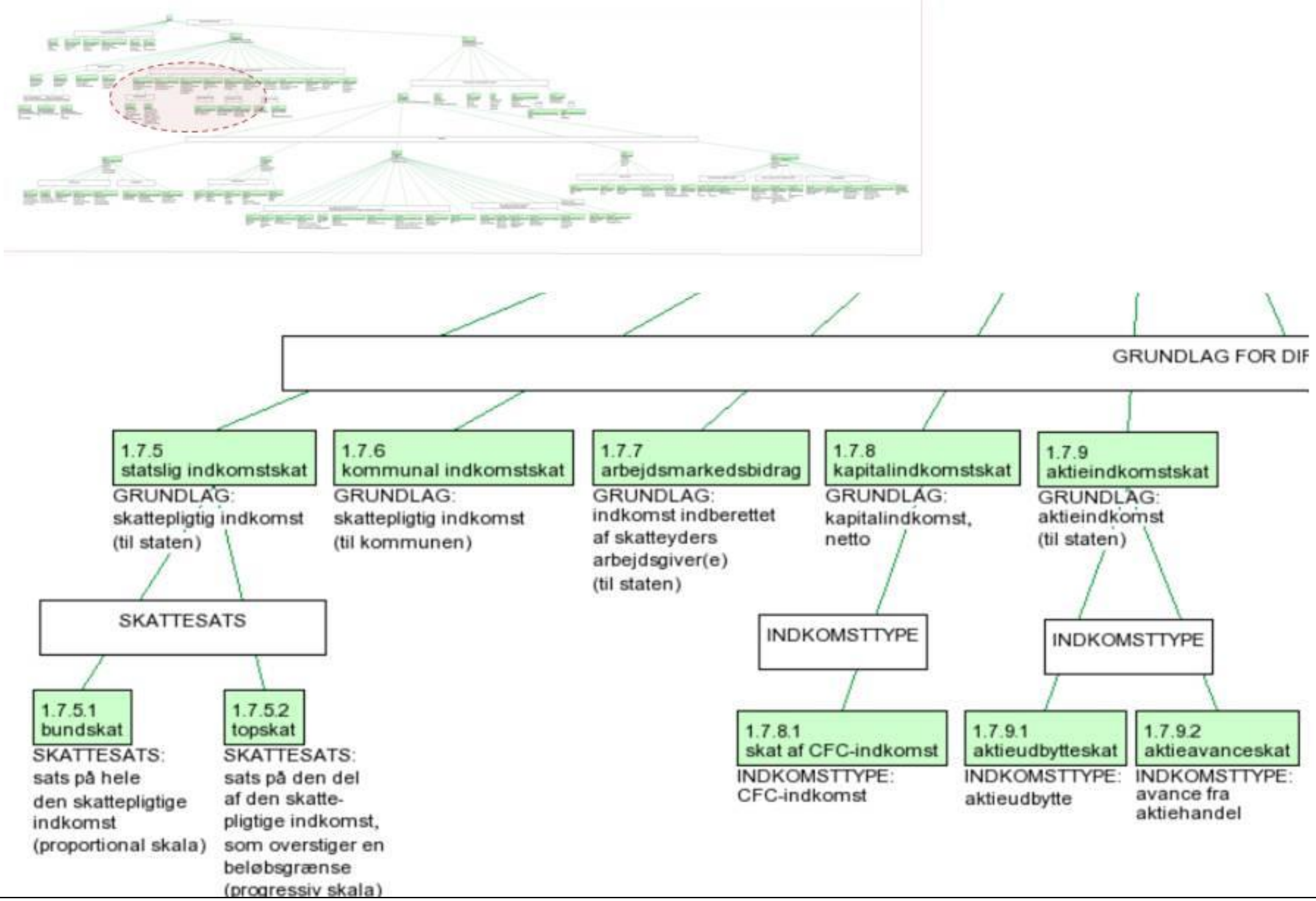


Figure II.5: Terminological ontology extract concerning the concept *direkte skat* (**1.7**), the subordinates **1.7.5-1.7.9** under the subdivision criteria: *GRUNDLAG FOR DIREKTE SKAT*, the sub-subordinates **1.7.5.1-1.7.5.2** under the subdivision criteria: *SKATTESATS*, and the sub-subordinates **1.7.8.1, 1.7.9.1-1.7.9.2** under the subdivision criteria: *INDKOMSTTYPE*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

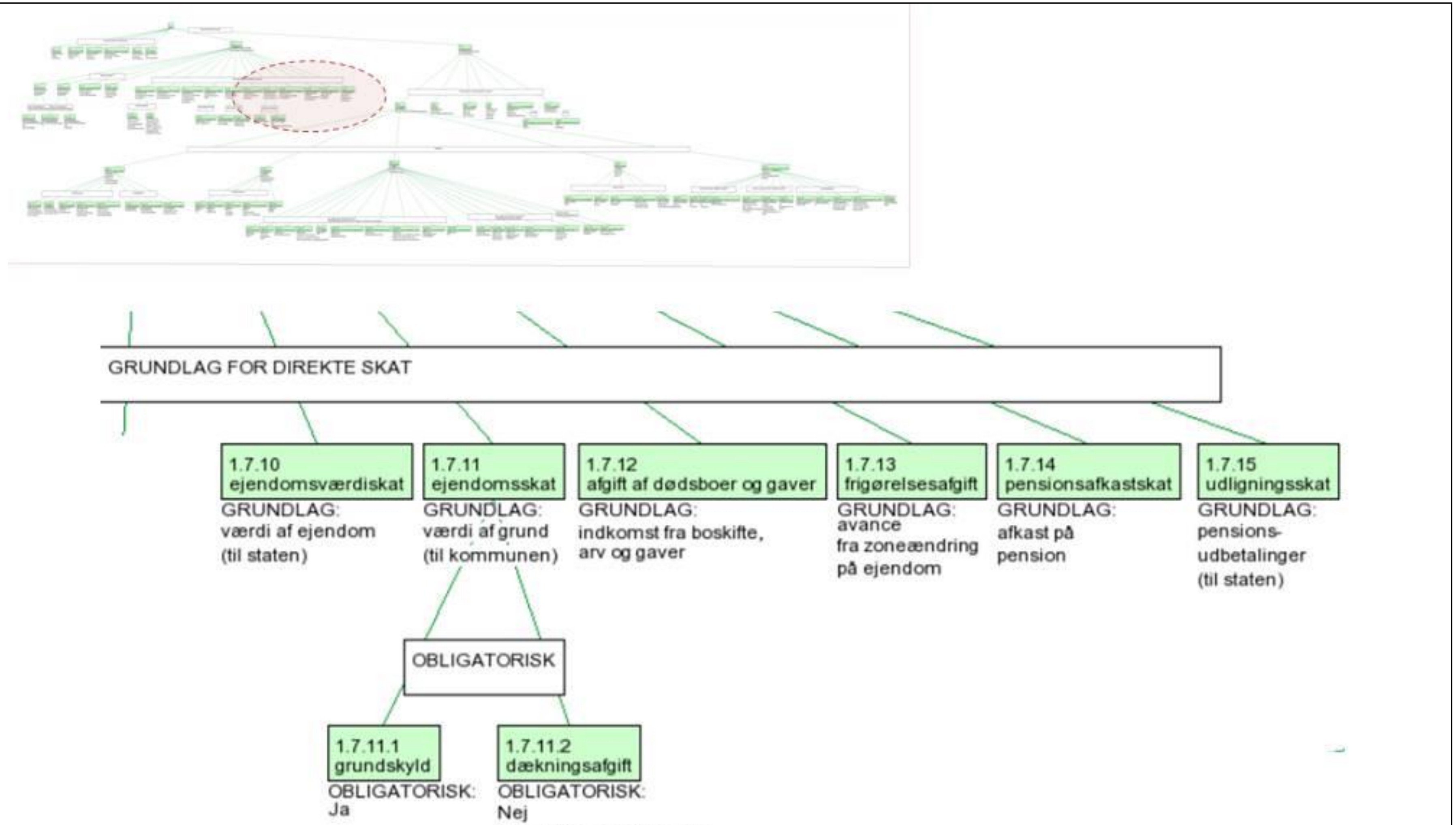


Figure II.6: Terminological ontology extract concerning the concept *direkte skat* (**1.7**), the subordinates **1.710.-1.7.15** under the subdivision criteria: *GRUNDLAG FOR DIREKTE SKAT*, and the sub-subordinates **1.7.11.1-1.7.11.2** under the subdivision criteria: *OBLIGATORISK*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

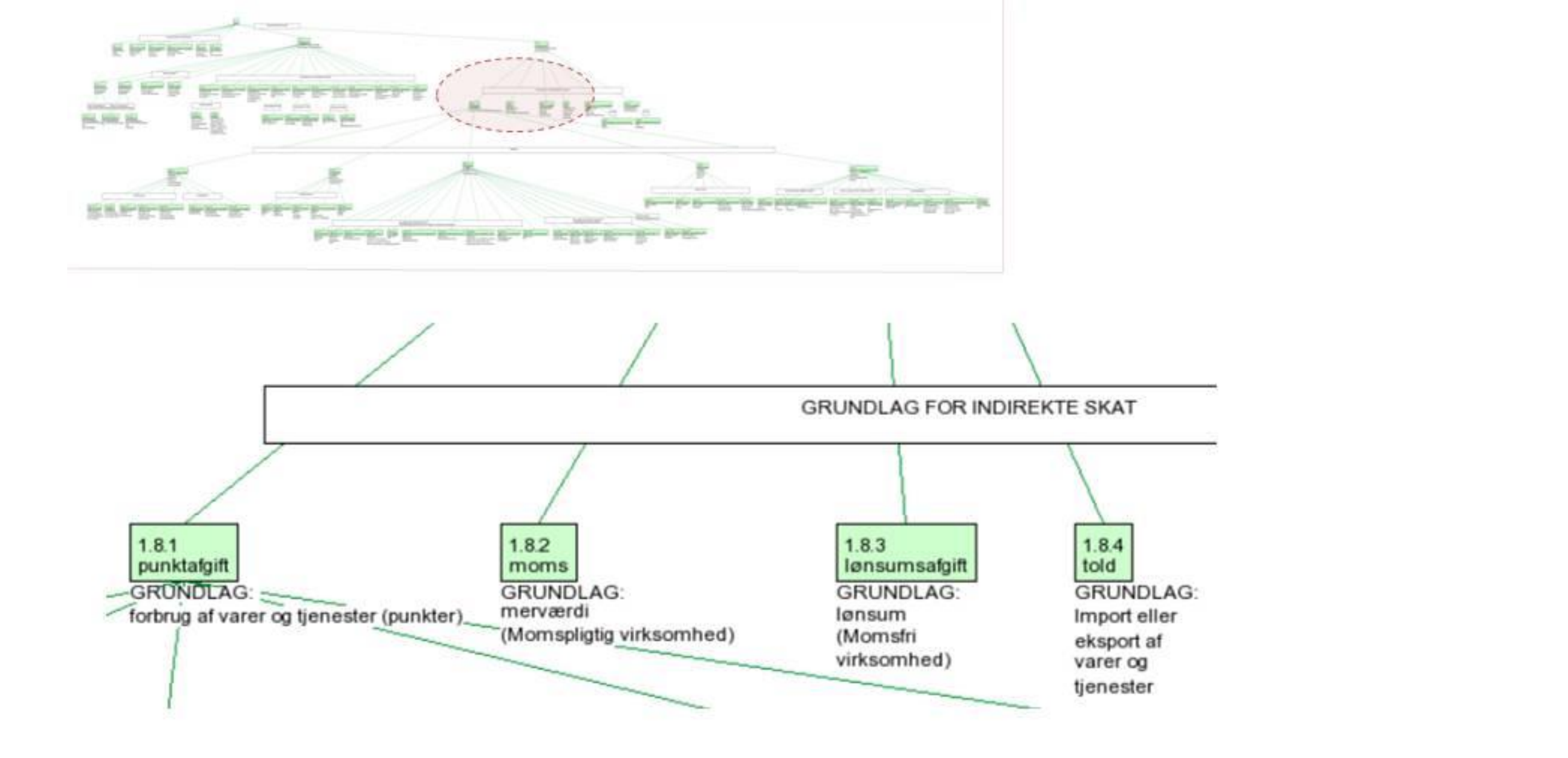


Figure II.7: Terminological ontology extract concerning the concept *indirekte skat* (**1.8**), and the subordinates **1.8.1-1.8.4** under the subdivision criteria: *GRUNDLAG FOR INDIREKTE SKAT*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

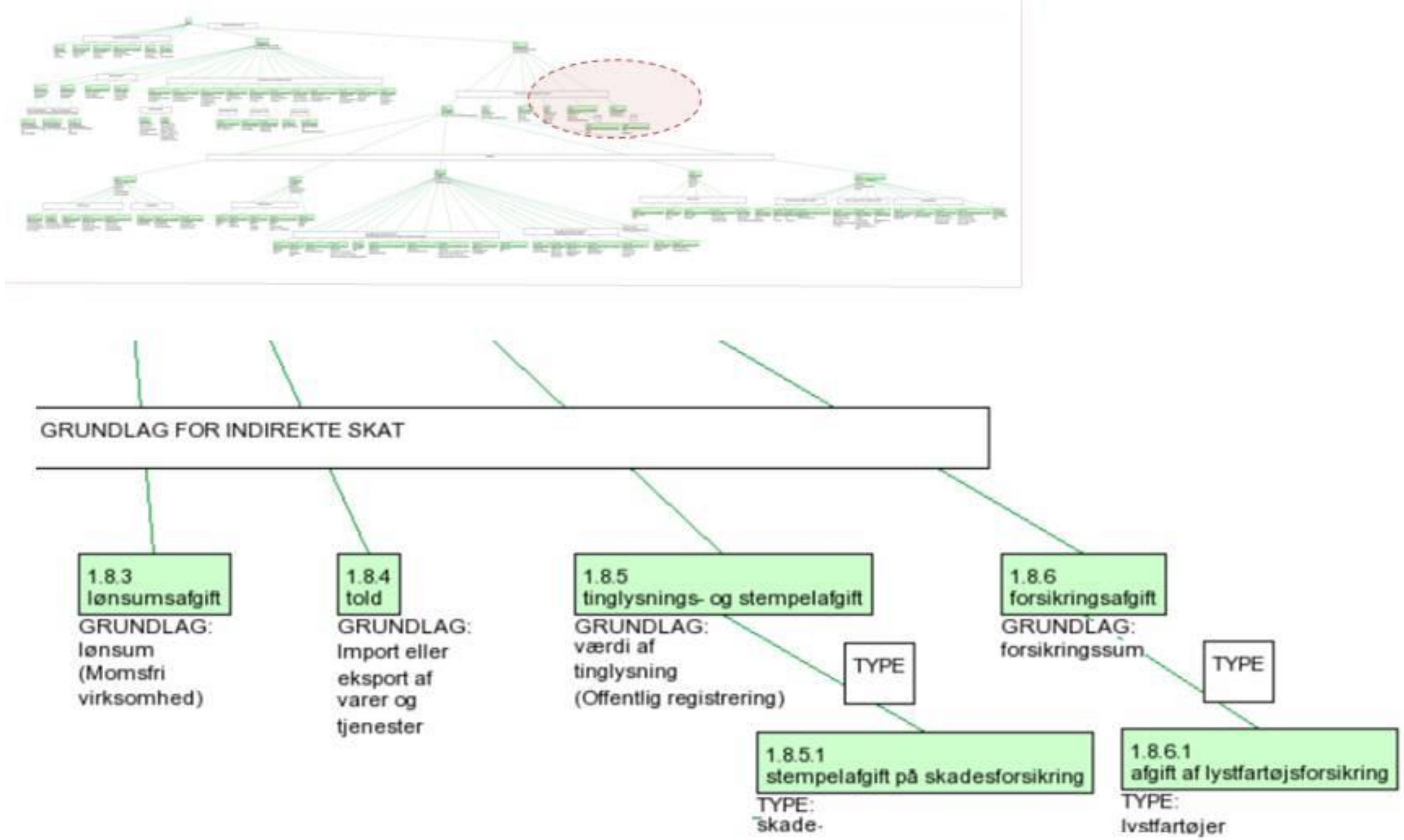


Figure II.8: Terminological ontology extract concerning the concept *indirekte skat* (**1.8**), the subordinates **1.8.3-1.8.6** under the subdivision criteria: *GRUNDLAG FOR INDIREKTE SKAT*, and the sub-subordinates **1.7.11.1-1.7.11.2** under the subdivision criteria: *TYPE*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

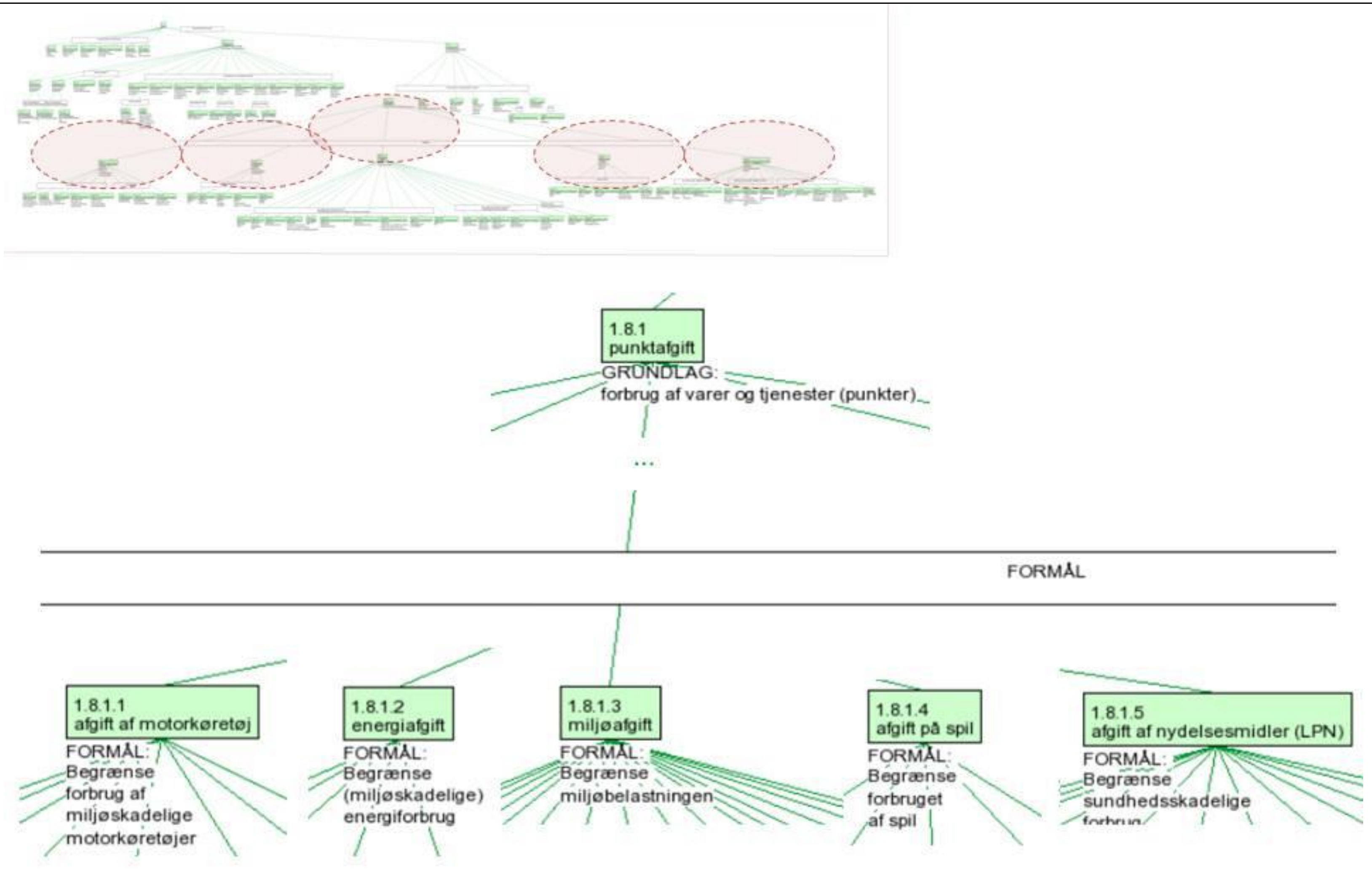


Figure II.9: Terminological ontology extract concerning the concept *punktafgift* (**1.8.1**), and the subordinates **1.8.1.1-1.8.1.5** under the subdivision criteria: *FORMÅL*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

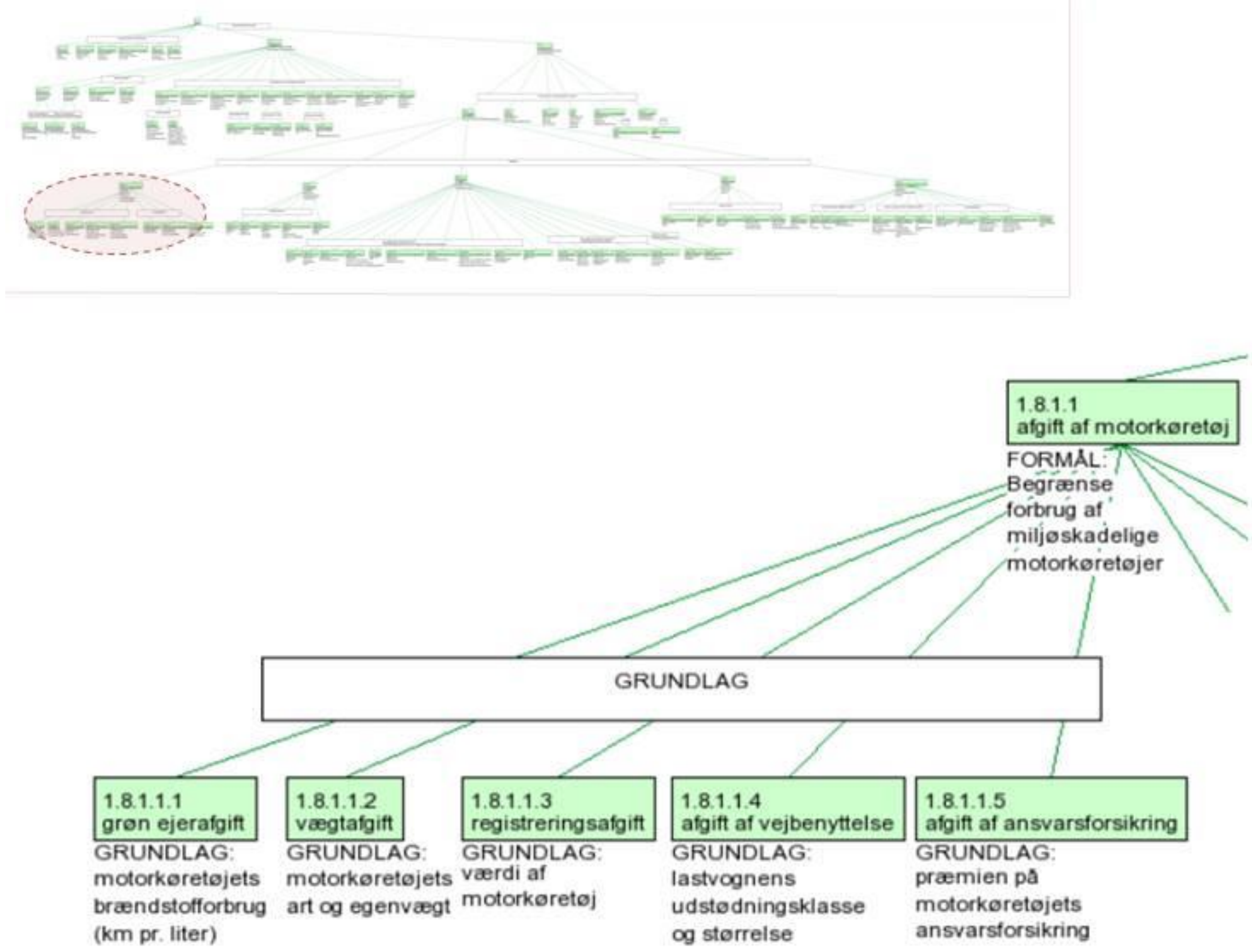


Figure II.10: Terminological ontology extract concerning the concept *afgift af motorkøretøj* (**1.8.1.1**), and the subordinates **1.8.1.1.1-1.8.1.1.5** under the subdivision criteria: *GRUNDLAG*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

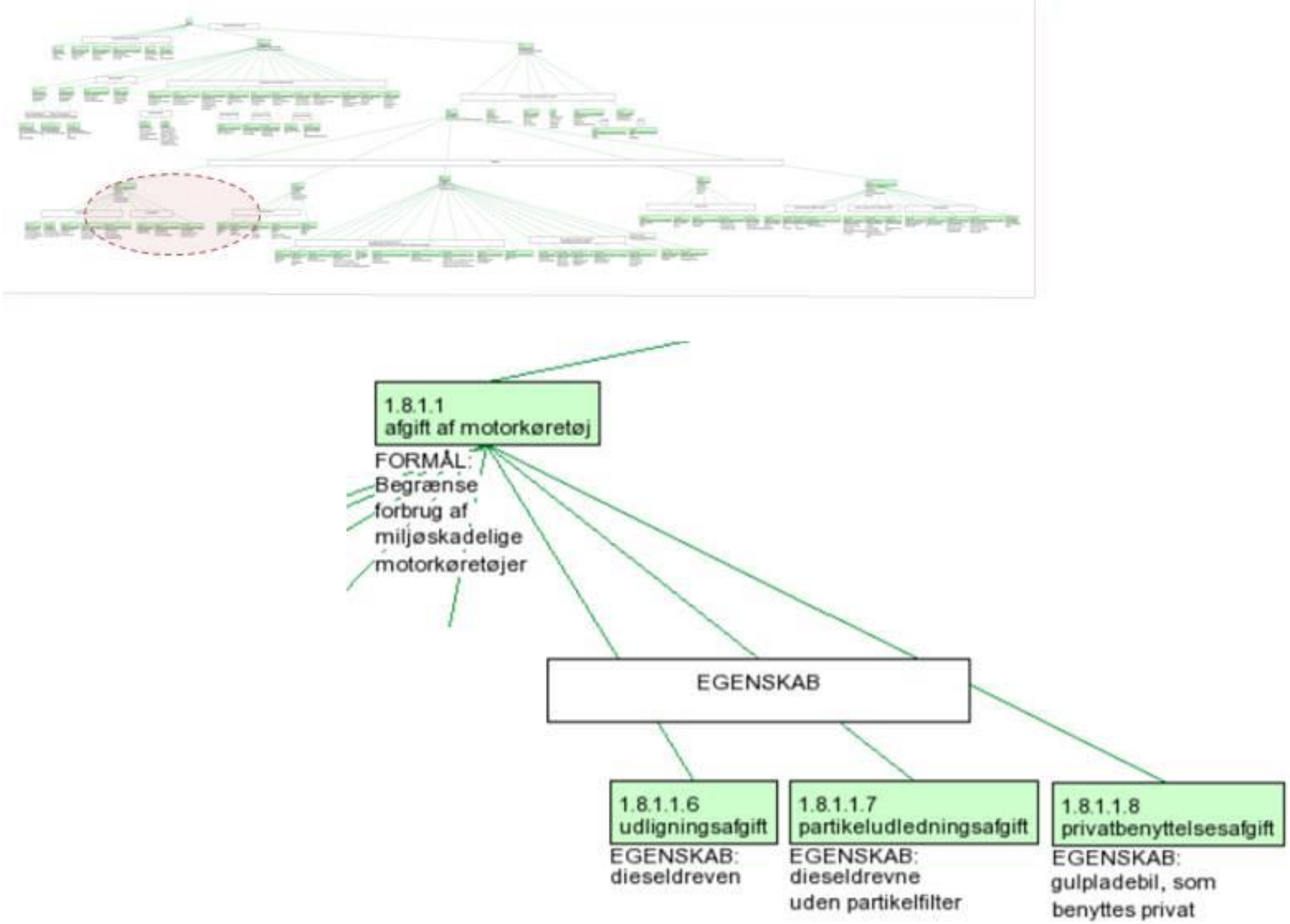


Figure II.11: Terminological ontology extract concerning the concept *afgift af motorkøretøj* (**1.8.1.1**), and the subordinates **1.8.1.1.6-1.8.1.1.8** under the subdivision criteria: *EGENSKAB*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

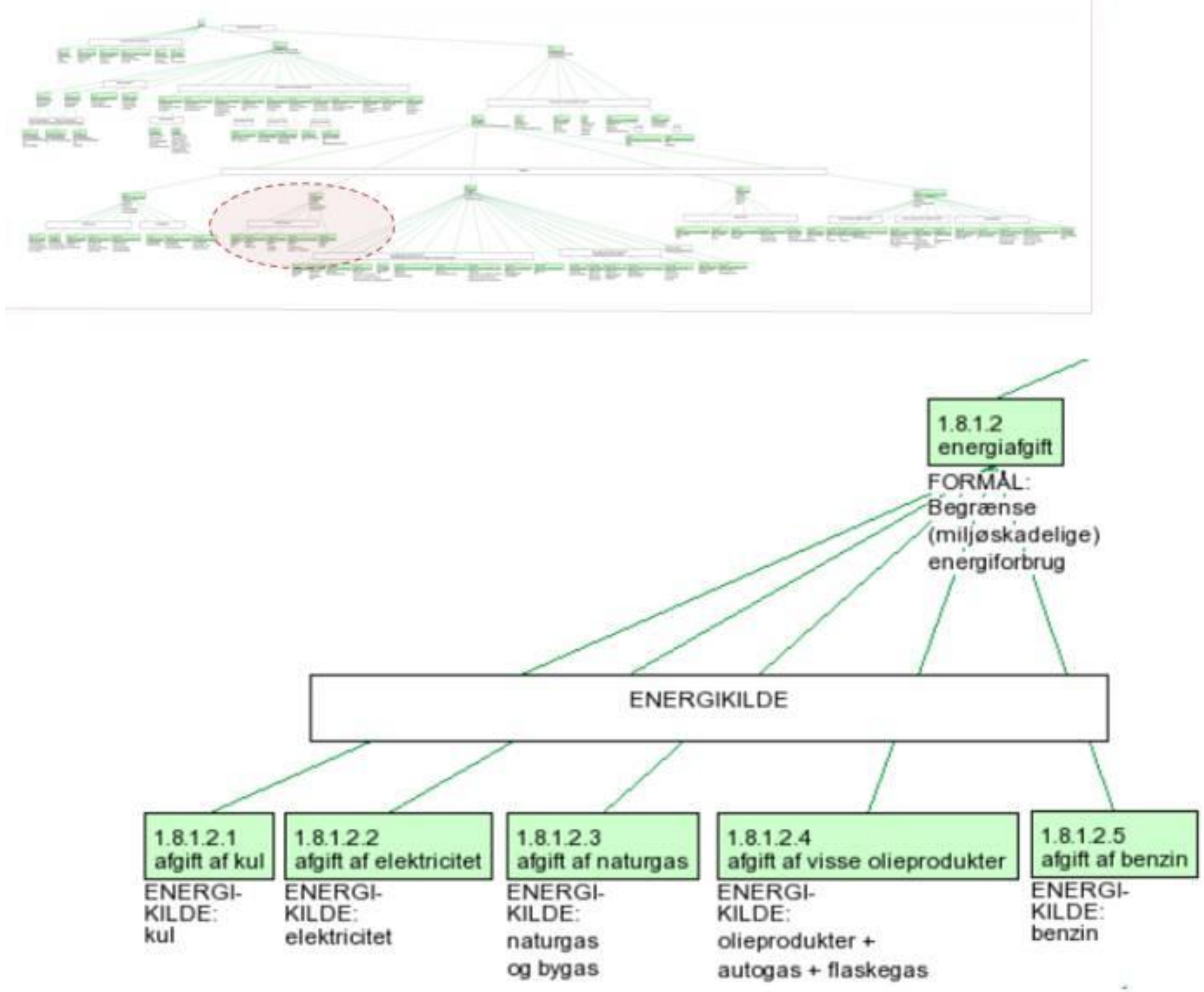


Figure II.12: Terminological ontology extract concerning the concept *energiafgift* (**1.8.1.2**), and the subordinates **1.8.1.2.1-1.8.1.2.5** under the subdivision criteria: *ENERGIKILDE*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

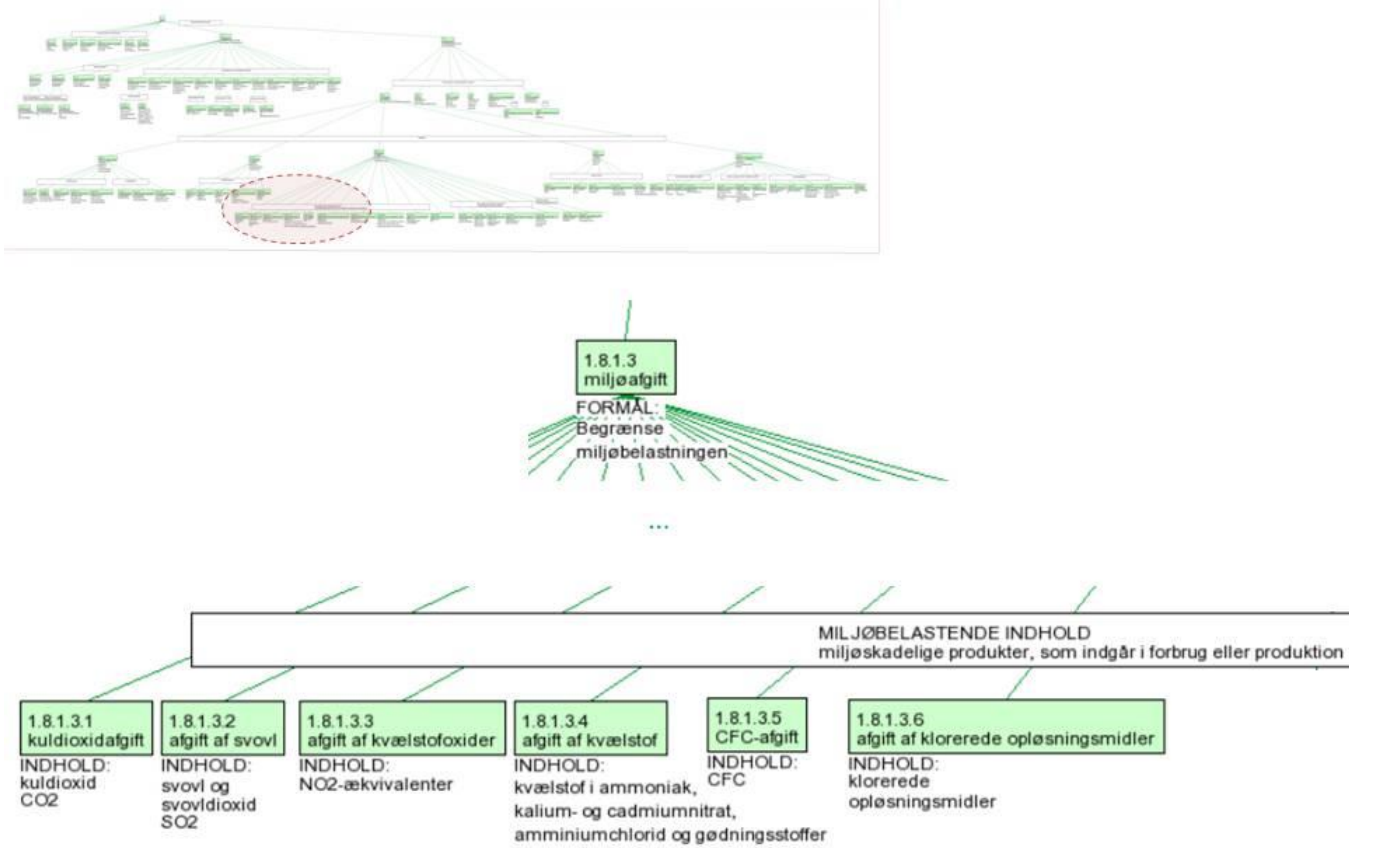


Figure II.13: Terminological ontology extract concerning the concept *miljøafgift* (**1.8.1.3**), and the subordinates **1.8.1.3.1-1.8.1.3.6** under the subdivision criteria: *MILJØBELASTENDE INDHOLD*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

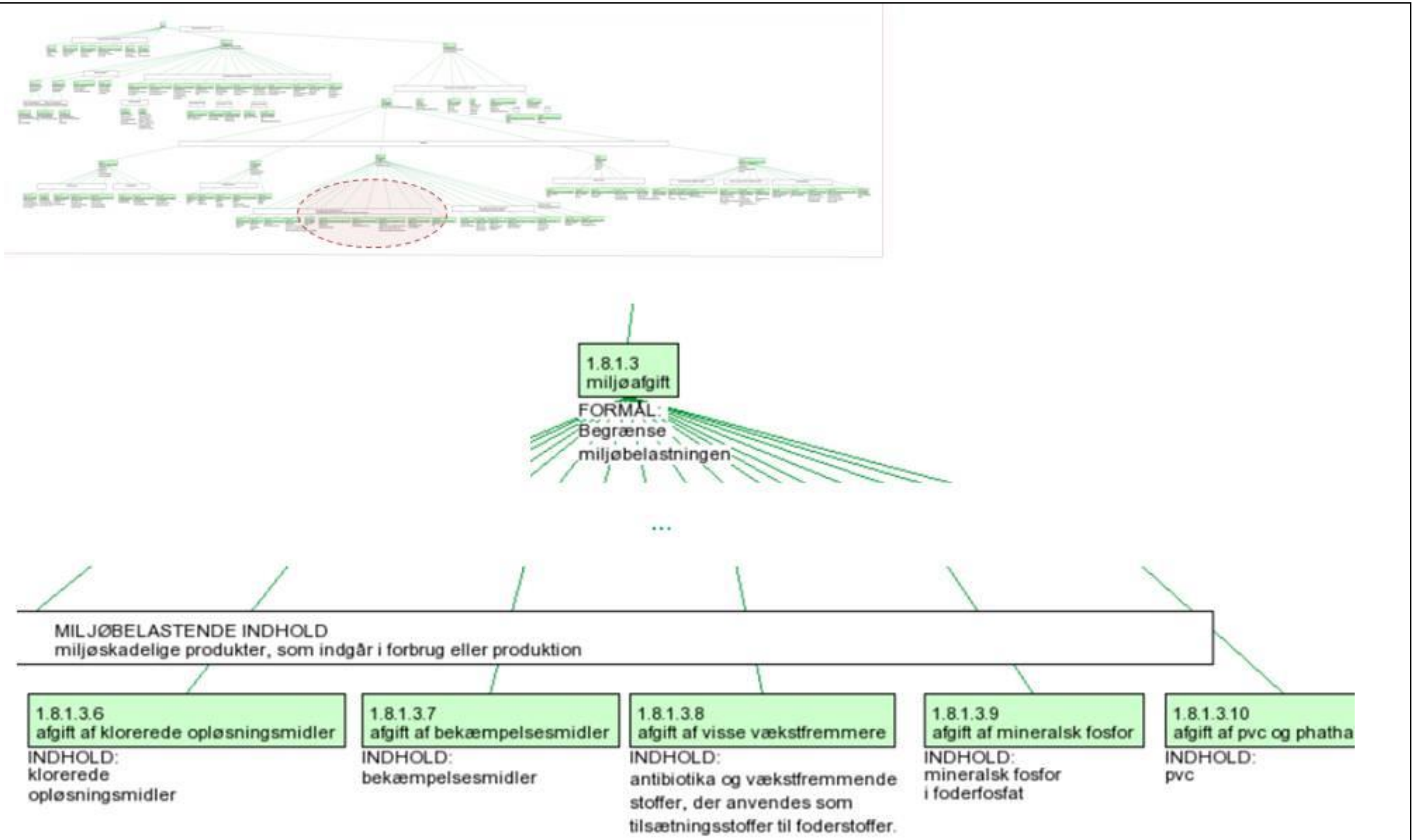


Figure II.14: Terminological ontology extract concerning the concept *miljøafgift* (**1.8.1.3**), and the subordinates **1.8.1.3.6-1.8.1.3.10** under the subdivision criteria: *MILJØBELASTENDE INDHOLD*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

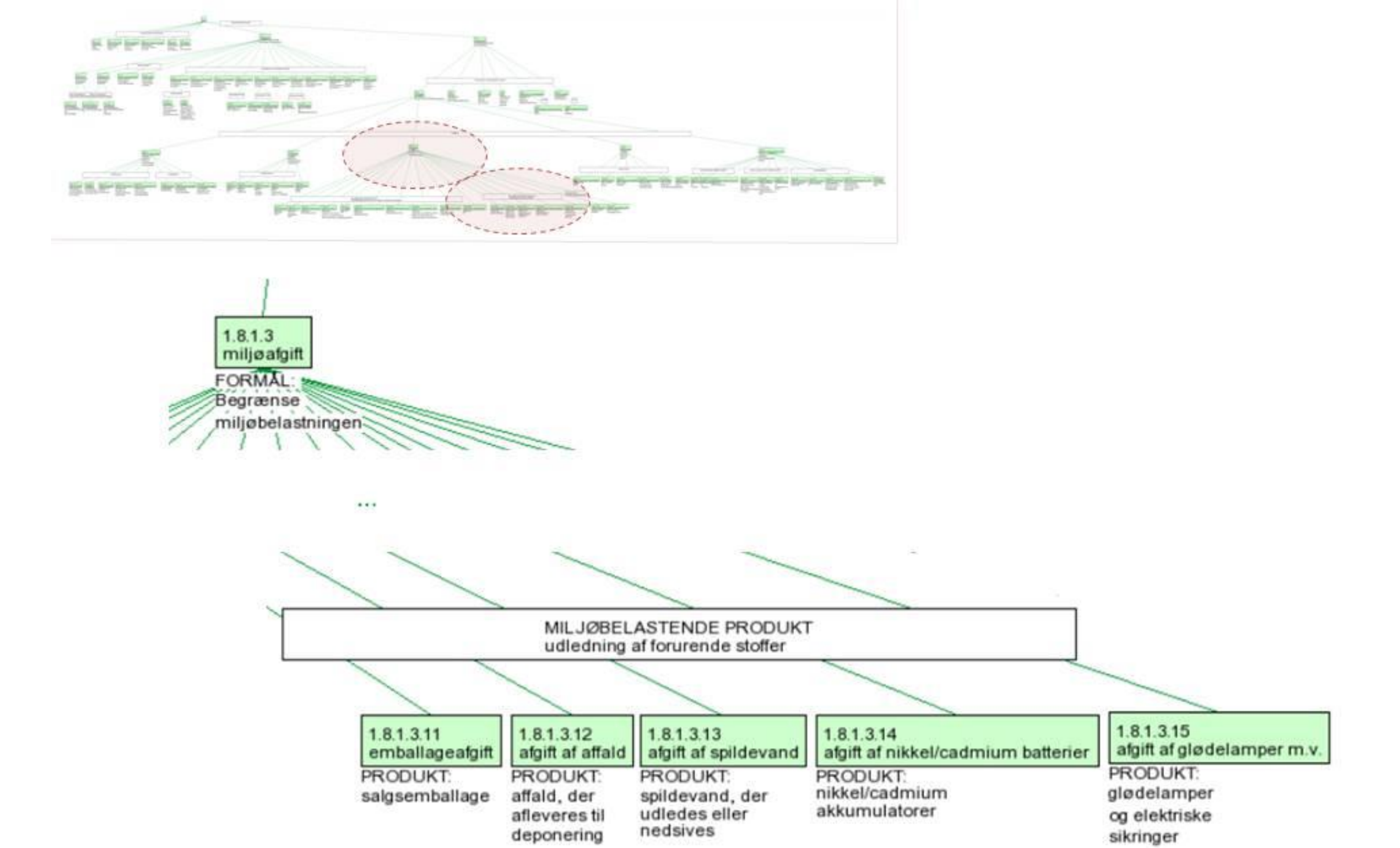


Figure II.15: Terminological ontology extract concerning the concept *miljøafgift* (**1.8.1.3**), and the subordinates **1.8.1.3.11-1.8.1.3.15** under the subdivision criteria: *MILJØBELASTENDE PRODUKT*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

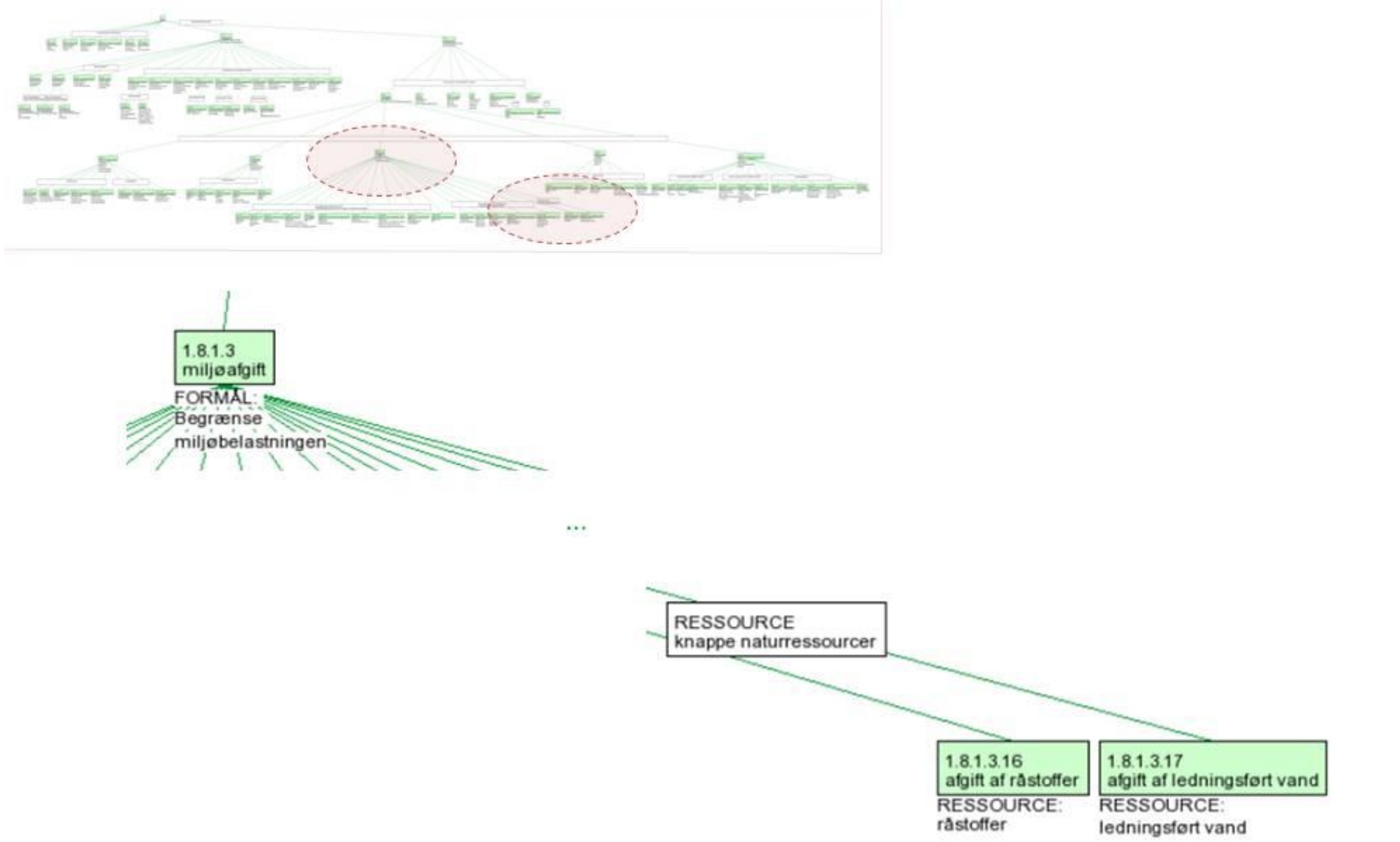


Figure II.16: Terminological ontology extract concerning the concept *miljøafgift* (**1.8.1.3**), and the subordinates **1.8.1.3.16-1.8.1.3.17** under the subdivision criteria: *RESSOURCE*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

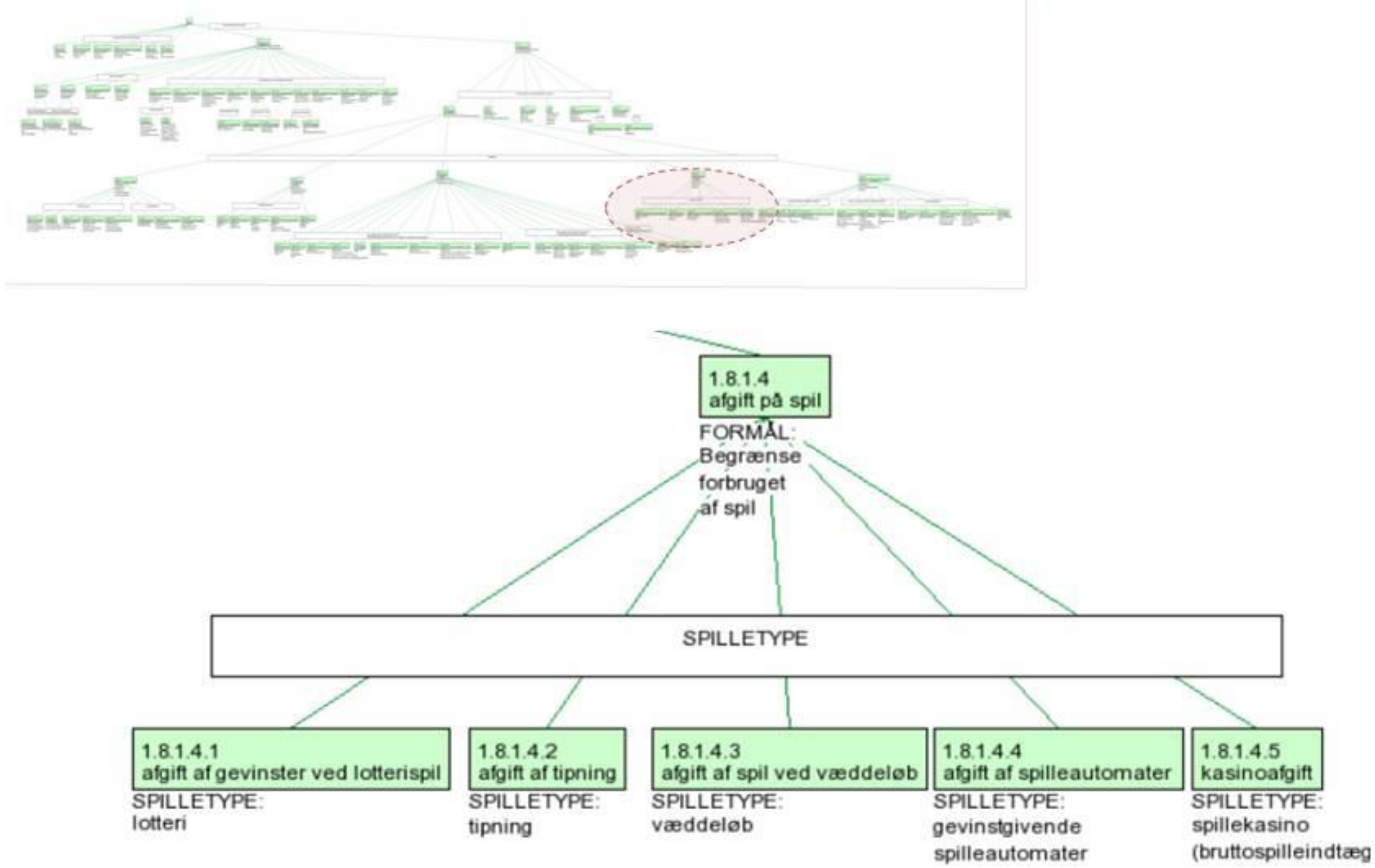


Figure II.17: Terminological ontology extract concerning the concept *afgift på spil* (**1.8.1.4**), and the subordinates **1.8.1.4.1-1.8.1.4.5** under the subdivision criteria: *SPILLETYPE*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

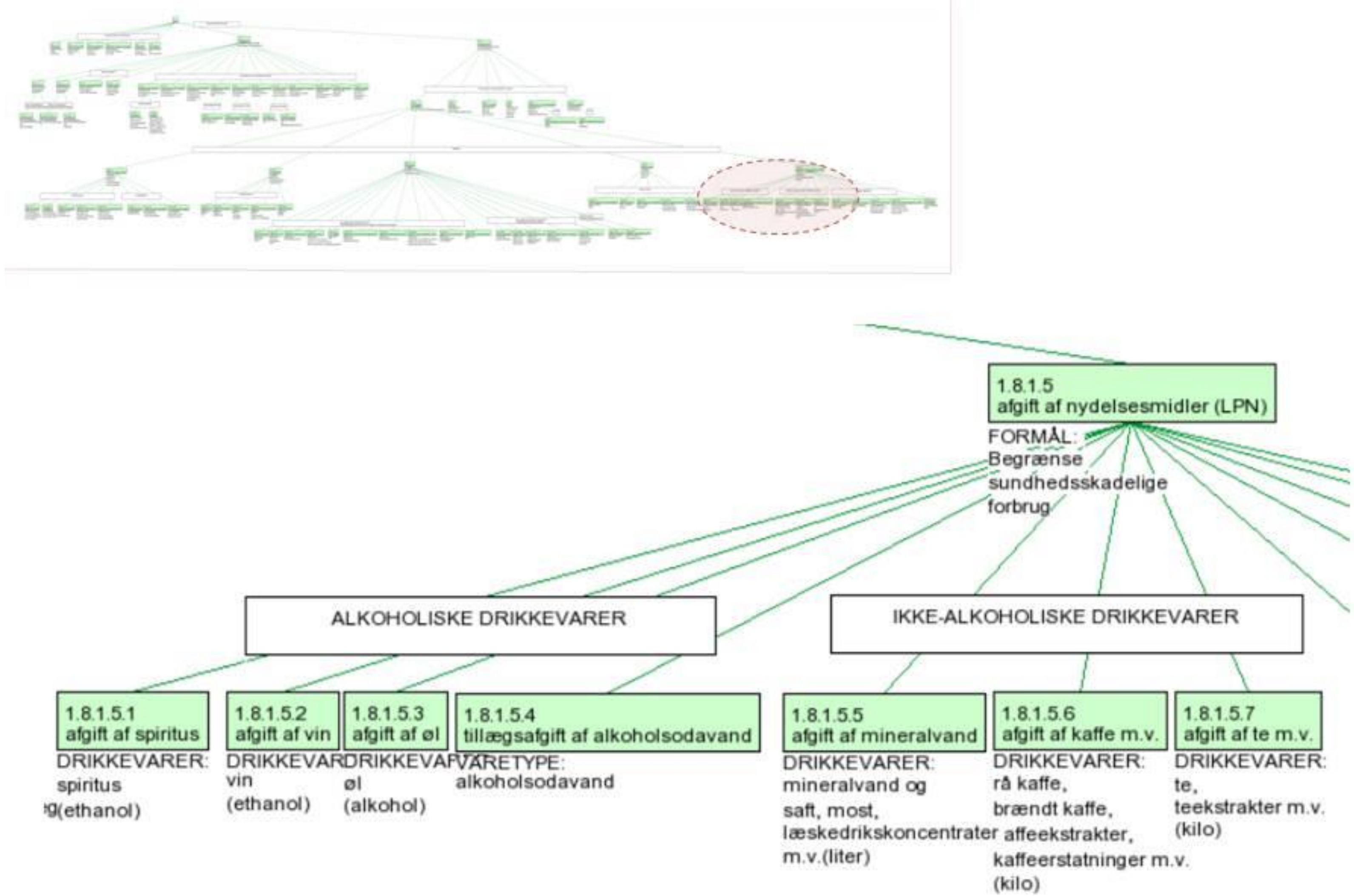


Figure II.18: Terminological ontology extract concerning the concept *afgift af nydelsesmidler* (**1.8.1.5**), the subordinates **1.8.1.5.1-1.8.1.5.4** under the subdivision criteria: *ALKOHOLISKE DRIKKEVARER*, and the subordinates **1.8.1.5.5-1.8.1.5.7** under the subdivision criteria: *IKKE-ALKOHOLISKE DRIKKEVARER*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

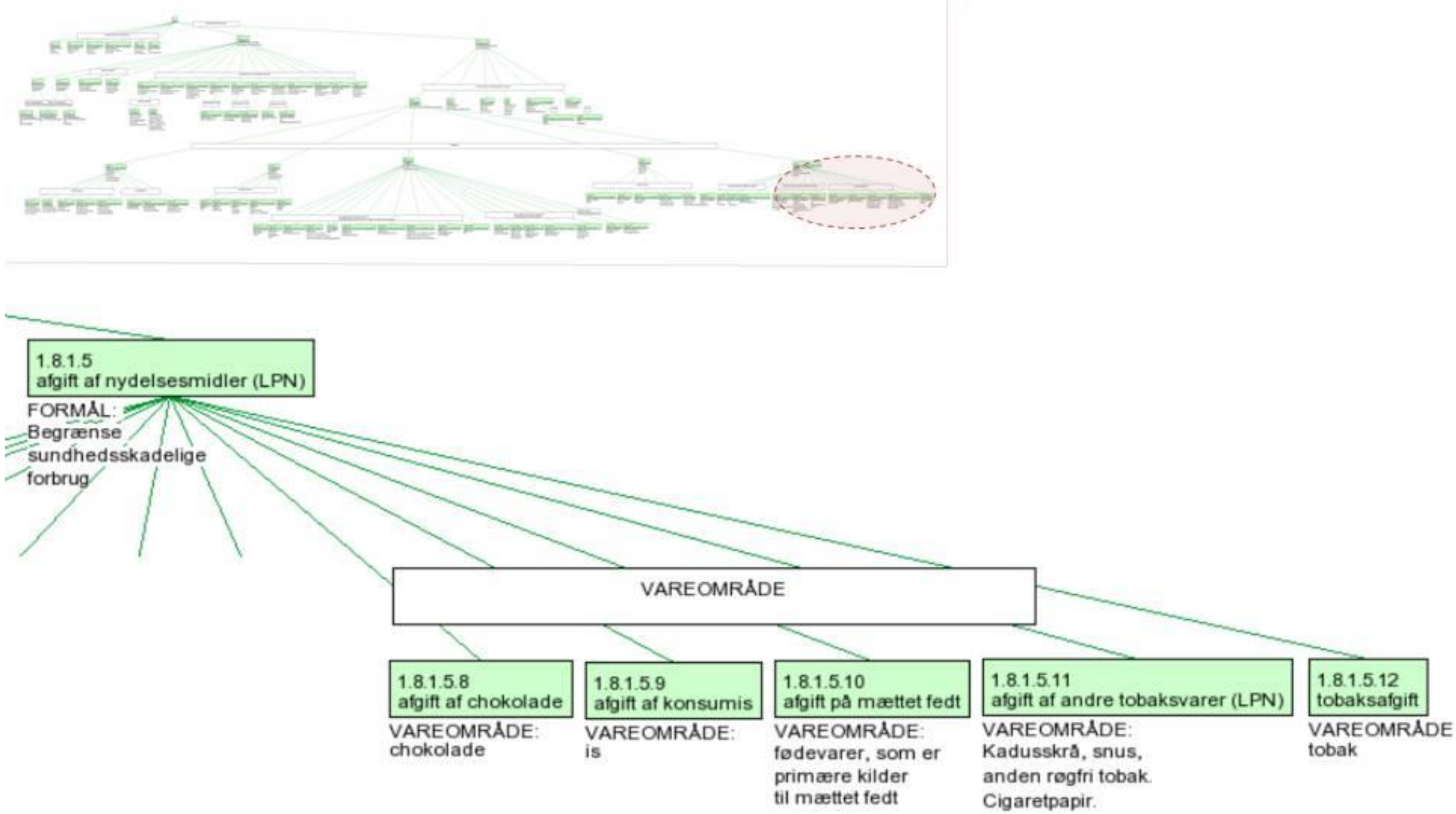


Figure II.19: Terminological ontology extract concerning the concept *afgift af nydelsesmidler* (**1.8.1.5**), and the subordinates **1.8.1.5.8-1.8.1.5.12** under the subdivision criteria: *VAREOMRÅDE*. The location of the extract in the entire terminological ontology is indicated by the overview in the top left corner. For translation into English see the systematic list in table I.2.

Appendix A:
Consent form and background questionnaire

Table A.1: Consent form. Participants signed the consent form upon briefing and prior to the experiments. Translation into English:

Section 1: ”Dear participant. You are about to participate in an experiment which is part of my PhD project ’User-oriented knowledge dissemination in the taxation domain’. The purpose of the project is to develop a user-interface for a data base (so-called term bank) providing the users’ need for knowledge. In particular, the question is whether information of the term bank can be disseminated to experts and non-experts, and whether the term bank may contribute to an increased understanding of technical concepts.”

Section 2: ”The experiment itself is conducted on a computer, which is connected to an eye-tracker registering your eye movements on-screen, while you will be asked to read a text, looking at grafic or schematic information about different technical terms as well as answering questions about the visualized technical terms. You will be giving your answers by pressing a key on the keyboard, and your response will be recorded and included in the subsequent analysis. It should be underlined that it is not a test of your knowledge in the specialist field, rather it is research about your behaviour as a user of a term bank.”

Section 3: ”Data collected during your participation is anonymized and cannot be connected to your person. This is the case for both the background questions that you will be answering prior to the experiment, your behaviour recorded by an eye-tracker and the keyboard during the experiment itself, as well as the answers that you give in the closing interview evaluating the experiment.”

Section 4: ”The results from the experiment will be published in my PhD dissertation and presented at the PhD defence. The results may also be part of scientific publications and presentations at scientific conferences, workshops, etc.”

Section 5: ”You have volunteered in participating. You may withdraw your consent at any time and without reason. If you decide to participate, I ask you to sign this consent form, which means that you approve of the use of data described in the previous sections. You have the right to time to think it over before you decide to sign.”

Section 6: ”Best regards; Louise Pram Nielsen; PhD Fellow at the Department of International Business Communication, CBS, Dalgas Have 15, 2000 Frederiksberg; Email: lpn.ibc@cbs.dk.”

Section 7: ”Date; Signature.”

Kære forsøgsdeltager.

Du skal nu deltage i et forsøg, som indgår i mit ph.d.-projekt ”Brugertilpasset vidensformidling på skatteområdet”. Projektet har til formål at udvikle en brugergrænseflade til en database (en såkaldt termbank), der imødekommer brugernes behov for viden. Spørgsmålet er især, om den information, der er i termbanken, kan formidles til eksperter og ikke-eksperter, og om termbanken kan give en øget forståelse af fagbegreber.

Selve eksperimentet foregår på en computer, som har en påmonteret eye-tracker, der registrerer dine øjnes bevægelser på skærmen, mens du bliver bedt om at læse en tekst, kigge på grafisk og skematisk information om forskellige fagudtryk samt svare på spørgsmål om de viste fagudtryk. Du afgiver dine svar ved at trykke på en tast på tastaturet, og din respons optages og indgår i den efterfølgende analyse. Det skal understreges, at det ikke er en test i din viden om fagområdet, men derimod en undersøgelse af din adfærd som bruger af en termbank.

Alle data, som er indsamlet i forbindelse med din deltagelse i forsøget, vil være anonymiseret og kan ikke kobles til din person. Det gælder både baggrundsspørgsmålene, som du skal besvare inden forsøget, din adfærd, som blev optaget af eye-trackeren og tastaturet under selve forsøget, samt de svar, du giver i det afsluttende interview, hvor vi evaluerer forsøget.

Resultaterne af forsøget vil blive publiceret i min ph.d.-afhandling og præsenteret i forbindelse med forsvaret af ph.d.-afhandlingen. Resultaterne vil også kunne indgå i videnskabelige publikationer og i indlæg ved videnskabelige konferencer, workshops mv.

Det er frivilligt, om du vil deltage i forsøget. Du kan når som helst og uden grund trække dit samtykke tilbage. Hvis du har besluttet dig for at deltage i forsøget, og du således tillader, at data benyttes som beskrevet ovenfor, vil jeg bede dig underskrive denne samtykkeerklæring. Husk, at du har ret til betænkningstid, før du beslutter, om du vil underskrive samtykkeerklæringen.

Med venlig hilsen
Louise Pram Nielsen
Ph.d.-studerende ved Department of International Business Communication, CBS, Dalgas Have 15, 2000 Frederiksberg. Email: lpn.ibc@cbs.dk

Dato: ______________________

Underskrift: ____

Table A.2: Background questionnaire. The background questions were answered by participants prior to the experiments. Translation into English:

Section 1: ”Gender. What is your sex? Female; Male.”

Section 2: ”Age. How old are you? Years.”

Section 3: ”Eye sight. Do you have normal or adjusted eyesight? Normal sight; adjusted: I use glasses/lenses to read or watch television; Adjusted eyesight: Other.”

Section 4: ”First language. Is Danish your first language? Yes; No.”

Section 5: ”English. How good is your English? Please, indicate your level on the line below (From very poor to very good).”

Section 6: ”Expertise. How big is your knowledge in the taxation domain? Please, indicate your level on the line below (From very small to very large)”

Section 7: ”What is your primary working areas currently? Please, indicate with one or several marks below. Working area: Administration, organisation and politics; Finance, insurance and banking; Research and teaching; HR and personnel; Skilled trade and service; IT and system development; Law and administration of justice; Consultant and counselling; Art, culture and sports; Management; Environment and nature; Natural sciences and technology; Sales and marketing; Social and health care; Language and communication; Transport and logistics; Finance and accounting. Please, describe your typical working tasks.”

Section 8: ”What is your highest level of completed education? Basic education; Skilled; Upper secondary; Short-term university education (1-2 years); Medium-term university education (3-4 years); Long-term university education (5-6 years or more). Please, state your title.”

Section 9: ”Information search. How good are you at searching information electronically. Please, indicate your level on the line below (From very poor to very good).”

Section 10: ”Please, state how often on a typical working week you use electronic search tools. You are welcome to add more, including printed works. **Vertical:** Google; Wikipedia; Gyldendal’s encyclopedia (*denstoredanske.dk*); EU’s multilingual term base (*iate.europa.eu*); Gyldendal’s dictionaries (*ordbog.gyldendal.dk*), Ordbogen.com’s dictionaries (*ordbogen.com*); Oxford dictionaries (*oxforddictionaries.com*); More: 1; 2; 3; 4; 5. **Horizontal:** 0 times, since I do not know the tool; 0 times, since I do not have access to the tool; 0 times, since I do not need the tool; 0 times, since I do not like the tool; 1-10 times; 10-50 times; 50-100 times; More than 100 times.”

Køn

Hvad er dit køn?

- ☐ Kvinde
- ☐ Mand

Alder

Hvor gammel er du?

___________ år

Syn

Har du normalt eller korrigeret syn?

- ☐ Normalt syn
- ☐ Korrigeret syn: Jeg bruger briller/kontaktlinser for at læse eller se fjernsyn
- ☐ Korrigeret syn: Andet

Modersmål

Er dansk dit modersmål?

- ☐ Ja
- ☐ Nej

Engelsk

Hvor god er du til engelsk? Sæt et kryds på nedenstående linje

Meget dårlig ———————————— Rigtig god

Ekspertise

Hvor stor en viden har du inden for skatteområdet? Sæt et kryds på nedenstående linje

Meget lidt ———————————— Rigtig meget

Hvad arbejder du <u>primært</u> med på nuværende tidspunkt? Sæt gerne flere krydser.

Arbejdsområde:

- ☐ Administration, organisation og politik
- ☐ Finans, forsikring og bank
- ☐ Forskning og undervisning
- ☐ HR og personale
- ☐ Håndværk og service
- ☐ IT og systemudvikling
- ☐ Jura og retsvæsen
- ☐ Konsulent og rådgivning
- ☐ Kunst, kultur og sport
- ☐ Ledelse
- ☐ Miljø og natur
- ☐ Naturvidenskab og teknik
- ☐ Salg og marketing
- ☐ Social og sundhed
- ☐ Sprog og kommunikation
- ☐ Transport og logistik
- ☐ Økonomi og regnskab

Typiske arbejdsopgaver :____

____

____

Hvad er den højeste uddannelse, du har gennemført:

- ☐ Grundskole
- ☐ Faglært
- ☐ Gymnasial
- ☐ Kort videregående (1-2 års varighed)
- ☐ Mellemlang videregående (3-4 års varighed)
- ☐ Lang videregående (5-6 års varighed eller længere)

Titel:____

Informationssøgning

Hvor god er du til at søge informationer elektronisk? Sæt et kryds på nedenstående linje

Meget dårlig — Rigtig god

Angiv hvor mange gange <u>på en typisk arbejdsuge</u> du bruger <u>elektroniske</u> søgeværktøjer. Tilføj gerne flere, også trykte publikationer.

	0 gange, da jeg ikke kender værktøjet	0 gange, da jeg ikke har adgang til værktøjet	0 gange, da jeg ikke har brug for værktøjet ugentligt	0 gange, da jeg ikke kan lide værktøjet	1-10 gange	10-50 gange	50-100 gange	Mere end 100 gange
Google	☐	☐	☐	☐	☐	☐	☐	☐
Wikipedia	☐	☐	☐	☐	☐	☐	☐	☐
Gyldendals åbne encyklopædi: denstoredanske.dk	☐	☐	☐	☐	☐	☐	☐	☐
EU's flersprogede termbase: iate.europa.eu	☐	☐	☐	☐	☐	☐	☐	☐
Gyldendals online ordbøger: ordbog.gyldendal.dk	☐	☐	☐	☐	☐	☐	☐	☐
Danmarks største online ordbog: ordbogen.com	☐	☐	☐	☐	☐	☐	☐	☐
Oxford Dictionaries: oxforddictionaries.com	☐	☐	☐	☐	☐	☐	☐	☐
Andre:								
1.					☐	☐	☐	☐
2.					☐	☐	☐	☐
3.					☐	☐	☐	☐
4.					☐	☐	☐	☐
5.					☐	☐	☐	☐

Appendix B:
Introduction, recalling and categorization

Table B.1: Introduction to concept clarification. The introduction was read by partipants upon completion of background questionnaire (see table A.2 above). Translation into English:

Section 1: ”Technical terms constitute words or expressions which are used in a specific specialist field with a precise, technical meaning. Technical terms are referred to as terms. Following that line of thought, a term bank is a data base containing structured information about technical terms. Information is visualized as diagrams showing the position of the technical term in the classification (see Figure 1), or as article entries providing detailed information about each individual term in one or more languages (see Figure 2).”

Section 2: ”Let us take a look at an example. We have encountered three concepts and their relations in a publication from Statistics Denmark. It says: 'The rules governing the personal tax calculation are to be found in the Personal tax law (Consolidated Act No. 959 of 19. September 2006 with later amendments). The tax calculation is based on a division of the taxable income into a personal income part and a capital income part.' By reading the Personal tax law we learn that the technical terms can be structured by means of a diagram.”

Figure 1: Title: Example of a diagram showing a chosen technical term from the Personal tax law. **Superordinate** (*skattepligtig indkomst*): 'taxable income'. **Instance** (*kapitalindkomst*) 'capital income' with subdivision criteria: INCOME SOURCE: interest, share price increase or dividends. **Sideordinate** (*personlig indkomst*): 'personal income' with subdivision criteria: INCOME SOURCE: personal income from wage, pension or transfer payments.

Section 3: ”The diagram shows an example of a so-called concept diagram, which contains one superordinate 'taxable income' and two subordinates 'personal income' and 'capital income' subdivided by the criteria 'income source'. In the concept diagram, the subdivision criteria 'income source' is visualized by a box covering both relations (the green lines) and by a caption below each concept.”

Section 4: ”From this system, it is now possible to formulate a definition for e.g. 'capital income' by stating the nearest superordinate as well as the value of the characteristic attribute subdividing 'capital income' from 'personal income', i.e. taxable income from interest, capital gains or dividends (net).”

Section 5: ”In addition, we can produce an article entry visualizing the collected information. The technical term (including its source) is stated together with the definition deduced from the concept diagram. Moreover, the example (including its source) that initiated the concept clarification is included. If the concept clarification is conducted in English and we believe that the English expression is a precise translation of the Danish technical term (equivalent), we may connect the two article entries.”

Figure 2: Title: Example of article entry covering capital income. **Row 1:** Danish: capital income. **Row 2:** Source for term: Taxes and Duties 2012 (Statistics Denmark). **Row 3:** Definition: taxable income from interest, capital gains or dividends (net). **Row 4:** Source: DanTermBank. **Row 5:** Comment: The tax calculation is based on a division of the taxable income into a personal income part and a capital income part. **Row 6:** Source: Taxes and Duties 2012 (Statistics Denmark).

Section 6: ”We saw that the definition of a technical term can be deduced from the concept diagram as well as the article entry. However, it should be noted that concept diagrams and article entries to a large extent provide different information. The strength of concept diagrams is the overview given by visualizing a concept's position i relation to related concepts, while the article entry gives more detailed information about each individual technical term, including translation into foreign languages.”

Et fagudtryk er ord eller udtryk som især bruges inden for et bestemt fagområde med en præcis, teknisk betydning. Fagudtryk kaldes også for termer, og en termbank er således en database, som indeholder strukturerede informationer om fagudtryk. Informationerne kan præsenteres som diagrammer, der viser fagudtrykkets placering i fagsystematikken (jf. figur 1), eller som artikler, der giver uddybende informationer om det enkelte fagudtryk på et eller flere sprog (jf. figur 2).

Lad os kigge på et eksempel. Vi har tre begreber og deres indbyrdes relationer, som vi har fundet i en publikation fra Danmarks Statistik. Her står: ”*Reglerne om personskatteberegningen findes i personskatteloven (lovbek. nr. 959 af 19. september 2006 med senere ændringer). Skatteberegningen er baseret på en opdeling af den skattepligtige indkomst i personlig indkomst og kapitalindkomst.”* Vi læser lidt af Personskatteloven og finder frem til, at vi kan strukturere fagudtrykkene i forhold til hinanden ved hjælp af følgende diagram:

Figur 1. Eksempel på et diagram, som viser udvalgte fagudtryk fra Personskatteloven

Diagrammet viser et udsnit af et såkaldt begrebssystem, som indeholder ét overbegreb ”skattepligtig indkomst” og to underbegreber ”personlig indkomst” og ”kapitalindkomst”, som kan adskilles efter kriteriet ”indtægtskilde”. I begrebssystemet illustreres kriteriet ”indtægtskilde” både med en kasse henover relationerne (de grønne streger) og med en tekstboks under begreberne.

Ud fra denne systematik, kan vi nu formulere en definition for f.eks. ”kapitalindkomst” ved at anføre det nærmeste overbegreb samt værdien på det karakteristiske træk, som adskiller ”kapitalindkomst” fra ”personlig indkomst”, dvs.:

skattepligtig indkomst fra renter, kursgevinst eller afkast (netto)

Vi kan desuden producere en artikel, som præsenterer de informationer, vi har indsamlet. Her angiver vi fagudtrykket (inklusive kilden) samt den definition, vi netop har udledt. Desuden kan vi angive det eksempel, som fik os i gang med begrebsafklaringen (inklusive kilden). Og hvis vi har foretaget samme øvelse på engelsk, og vi vurderer, at det danske fagudtryk er en præcis oversættelse til det engelske (ækvivalent), kobler vi artiklerne sammen:

Figur 2. Eksempel på artikel over kapitalindkomst

Dansk: Kilde til term:	**kapitalindkomst** Skatter og afgifter 2012 (Danmarks Statistik)
Definition: Kilde:	skattepligtig indkomst fra renter, kursgevinst eller afkast (netto). DanTermBank
Kommentar: Kilde:	Personskatteberegningen er baseret på en opdeling af den skattepligtige indkomst i personlig indkomst og kapitalindkomst. Skatter og afgifter 2012 (Danmarks Statistik
English: Source for term:	**capital income** Taxation trends in the EU 2012 (Eurostat)
Definition: Source:	positive net capital income subject to taxation DanTermBank

Vi så ovenfor, at definitionen på et fagudtryk både vil kunne udledes af diagrammet og genfindes i artiklen. Det skal dog bemærkes, at diagrammer og artikler i høj grad opfylder forskellige informationsbehov. Styrken ved diagrammet er således, at det giver et godt overblik ved at illustrere et begrebs placering i forhold til relaterede begreber, mens artiklen til gengæld kan give mere uddybende informationer om det enkelte fagudtryk, herunder oversættelse til fremmedsprog.

Table B.2: Recalling taxation terms. The task was conducted by participants after the introduction to concept clarification (see table B.1 above). Translation into English: ”Please, write the first ten technical terms from the taxation field that comes to your mind (exclude the ones you read above). If you are not able to come up with that many, please, feel free to state fewer.”
Skriv de første ti fagudtryk på skatteområdet, du kan komme i tanke om (se bort fra dem, du så ovenfor). Hvis du ikke kan komme i tanke om så mange, må du gerne nøjes med færre. 1. __________ 2. __________ 3. __________ 4. __________ 5. __________ 6. __________ 7. __________ 8. __________ 9. __________ 10. __________

Table B.3: Categorization of terms and pseudo-terms. The task was conducted by participants after the recalling of taxation terms (see table B.2 above). Translation into English: ”For each of the following expression, please evaluate whether the expression exists (i.e. it is a technical term, which is being or has been used in the taxation field) or not. If you believe that the expression exists, please, indicate whether you know the meaning of the expression (in which case you are asked to describe the meaning in a few words). **1** (*ejendomsskat*): ’land tax’; **2** (*mellemskat*): ’middle-bracket tax’; **3** (*mindreværdiafgift*): ’value-subtracted tax’; **4** (*skatteindtægt*): ’tax receipts; **5** (*forskudsfradrag*): ’provisional deduction’; **6** (*forbrugsbegrænsende afgift*): ’consumption limiting duty’; **7** (*kaffemoms*): ’coffee VAT’; **8** (*vindafgift*): ’wind duty’; **9** (*energiafgift*): ’energy duty’; **10** (*grøn afgift*): ’green tax’; **11** (*C-skat*): ’C-tax’; **12** (*skatteprovenu*): ’tax revenue’; **13** (*fiskal skat*): ’fiscal tax’; **14** (*virksomhedsskat*)**:** ’corporation tax’; **15** (*personlig indkomstskat*): ’personal income tax’. **Categories:** ”No; the expression does not exist”; ”Yes, the expression exists, but I do not know the meaning”; ”Yes, the expression exists, and it means.”

Du skal nu for hvert af de nedenstående udtryk vurdere, om udtrykket eksisterer (dvs. om det er et fagudtryk, som benyttes eller er blevet benyttet på skatteområdet). Hvis du mener, at udtrykket eksisterer, skal du angive, om du kender betydningen af udtrykket (og i så fald skal du med få ord beskrive, hvad det betyder).

1. ejendomsskat

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

____

2. mellemskat

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

____

3. mindreværdiafgift

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

____

4. skatteindtægt

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

____

5. forskudsfradrag

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

____

6. forbrugsbegrænsende afgift

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

____

7. kaffemoms

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

____

8. vindafgift

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

____

9. energiafgift

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

____

10. grøn afgift

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

___

11. C-skat

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

___

12. skatteprovenu

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

___

13. fiskal skat

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

___

14. virksomhedsskat

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

___

15. personlig indkomstskat

- ☐ Nej, udtrykket eksisterer ikke
- ☐ Ja, udtrykket eksisterer, men jeg kender ikke betydningen af udtrykket.
- ☐ Ja, udtrykket eksisterer, og det betyder:

___

Appendix C:
Scenario and reading task

Velkommen.

Du deltager i et forsøg, hvor du vil få to typer opgaver: Du begynder med at læse en tekst og bagefter skal du svare på en række spørgsmål om fagudtryk fra skatteområdet.

Resultatet af forsøget vil gøre os klogere på brugeradfærd i forbindelse med søgning i databaser. Det er vigtigt, at du gør dit bedste. Du skal derfor gøre dig umage med at svare korrekt og præcist, så hurtigt som muligt.

Forsøget varer omkring 20 minutter.
Når du har løst alle opgaverne, tager vi en snak om din opfattelse af forsøget.

God fornøjelse!

Tryk på mellemrum, når du er klar til at fortsætte.

Figure C.1: Introduction to the experiment. Translation into English: ”Welcome. You are participating in an experiment, where you will get two types of tasks: To begin with you will read a text and then you are asked to answer a number of questions concerning expressions from the taxation domain. The results will add to our knowledge about user behaviour in searching databases. It is important that you try doing your best. We urge you to answer as correctly and accurately as possible, as quick as possible. The experiment takes about 20 minutes. When you are finished, we will discuss your view of the experiment. Have fun! Please, press space when you are ready to continue.”

Figure C.2: Description of the scenario and instruction to the reading task. Translation into English: ”Imagine that you work on a newspaper writing articles about the European economy. Today you are working on an article about the Danish taxation system since Statistics Denmark has published new figures. Therefore, you begin by reading an excerpt from ’Taxes and Duties,’ which has just been published by Statistics Denmark.”

"Skatteindtægterne er den vigtigste kilde til finansiering af de offentlige udgifter, og der er derfor en tæt sammenhæng mellem det samlede skatteprovenu og niveauet for de offentlige udgifter og overskuddet på de offentlige finanser.

Nogle skatter og afgifter er ikke rene fiskale skatter, dvs. at de ikke alene er begrundet med behovet for finansiering af de offentlige udgifter, men tilsigter primært at påvirke borgernes eller erhvervslivets adfærd. Det gælder fx afgifterne på energi og forurening, som fordyrer ydelserne og dermed begrænser forbruget af dem. Også afgifterne på alkohol og tobak nævnes ofte som sådanne forbrugsbegrænsende afgifter.

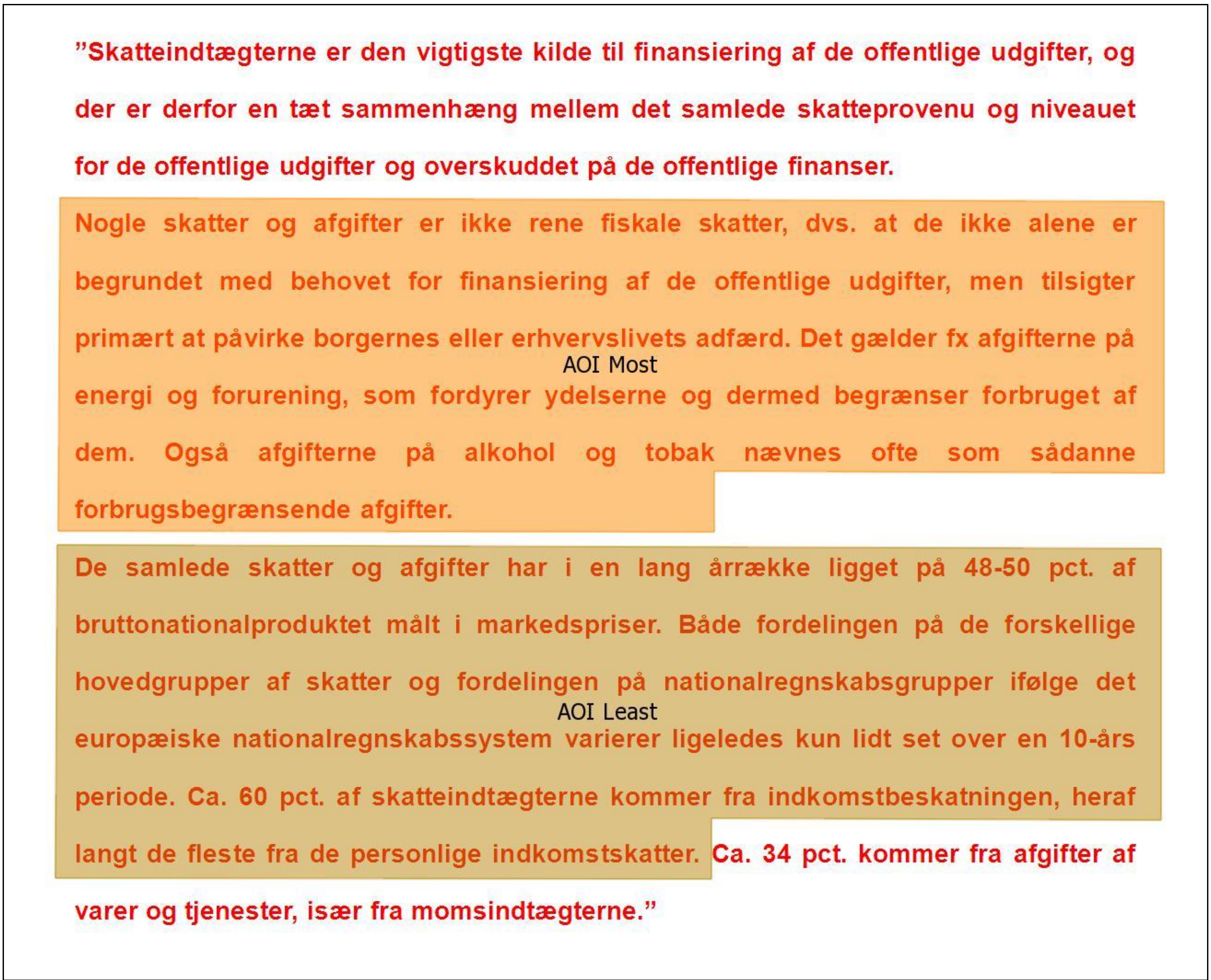
De samlede skatter og afgifter har i en lang årrække ligget på 48-50 pct. af bruttonationalproduktet målt i markedspriser. Både fordelingen på de forskellige hovedgrupper af skatter og fordelingen på nationalregnskabsgrupper ifølge det europæiske nationalregnskabssystem varierer ligeledes kun lidt set over en 10-års periode. Ca. 60 pct. af skatteindtægterne kommer fra indkomstbeskatningen, heraf langt de fleste fra de personlige indkomstskatter. Ca. 34 pct. kommer fra afgifter af varer og tjenester, især fra momsindtægterne."

Figure C.3: Reading task. The specialized text comprising the reading task was presented in font Arial, colour red, size 16 and double-spaced. Translation into English:

Section 1: "The tax receipts is the primary source of income financing the public spending, which provides a close relation between the total tax revenue and the level of public spending and the surplus on public finances."

Section 2: "Some types of taxes and duties do not constitute pure fiscal taxes, i.e. they are not only a matter of the need for financing the public spending, but aim at influencing behaviour of citizens and companies. This applies to e.g. duties on energy and pollution, which increases the price of the products, decreasing their consumption. Also, the duties on alcohol and tobacco are often considered to be corresponding consumption-limiting duties."

Section 3: "For many years, total taxes and duties have amounted to 48-50 per cent of the gross domestic product in market prices. The distribution across different main groups of taxes and the distribution across national-accounting groups according to the European national accounting system varies only to a minor degree when considered over a ten-year period. Approximately 60 per cent of the tax receipts comes from the income taxation, primarily personal income taxes. Almost 34 per cent are from duties on goods and services, in particular, income from value-added taxes."



"Skatteindtægterne er den vigtigste kilde til finansiering af de offentlige udgifter, og der er derfor en tæt sammenhæng mellem det samlede skatteprovenu og niveauet for de offentlige udgifter og overskuddet på de offentlige finanser.

Nogle skatter og afgifter er ikke rene fiskale skatter, dvs. at de ikke alene er begrundet med behovet for finansiering af de offentlige udgifter, men tilsigter primært at påvirke borgernes eller erhvervslivets adfærd. Det gælder fx afgifterne på energi og forurening, som fordyrer ydelserne og dermed begrænser forbruget af dem. Også afgifterne på alkohol og tobak nævnes ofte som sådanne forbrugsbegrænsende afgifter.

De samlede skatter og afgifter har i en lang årrække ligget på 48-50 pct. af bruttonationalproduktet målt i markedspriser. Både fordelingen på de forskellige hovedgrupper af skatter og fordelingen på nationalregnskabsgrupper ifølge det europæiske nationalregnskabssystem varierer ligeledes kun lidt set over en 10-års periode. Ca. 60 pct. af skatteindtægterne kommer fra indkomstbeskatningen, heraf langt de fleste fra de personlige indkomstskatter. Ca. 34 pct. kommer fra afgifter af varer og tjenester, især fra momsindtægterne."

Figure C.4: Reading task marked with equally sized areas-of-interest (AOIs): The upper AOI is the most complex (AOI Most), while the lower is the least complex (AOI Least).

Appendix D:
Dual-entry-mode experiment

Skatteområdet kan være et kompliceret område at formidle. Det ved du fra dit arbejde med europæisk økonomi på avisen. Du har derfor besluttet at forsyne din artikel med en boks, som skal give dine læsere forklaringer på betydningen af udvalgte fagudtryk fra din artikel.

Du får nu vist otte forskellige fagudtryk fra skatteområdet,
og der stilles seks spørgsmål til hvert fagudtryk dvs. i alt 48 spørgsmål.
Hvis du mener, at den første svarmulighed er korrekt, vælger du ”1”.
Hvis du mener, at den anden svarmulighed er korrekt, vælger du ”2”.
Og hvis du mener, at den tredje svarmulighed er korrekt, vælger du ”3”.

Du kan finde svaret ved at kigge på teksten og grafikken under spørgsmålet.
Der er kun et rigtig svar til hvert spørgsmål. Når du har svaret, vil du få at vide,
om du har svaret rigtigt eller forkert.

Figure D.1: Description of scenario and instruction to dual-mode experiment. The scenario description is continued from the reading task (see figure C.2 above) and instruction to answering the multiple-choice questions of the dual-mode experiment is outlined. Translation into English:
Section 1: ”The taxation domain may be a complex area to disseminate. You know this from your work with European economy on the newspaper. Therefore, you have decided to provide your article with a box giving your target readers explanations to the meaning of chosen technical terms occuring in your article.”
Section 2: ”You are presented with eight different technical terms from the taxation field and you will get six questions to each, i.e. 48 questions in total. If you believe that the first available answer is correct, please press ’1’. If you believe that the second available answer is correct, please press ’2’. And if you believe that the third available answer is correct, please press ’3’. You will be able to locate the answer by looking at the text and graphics below each question. There is only one correct answer to each question. Once you have answered, you are informed whether the answer was correct or incorrect.”

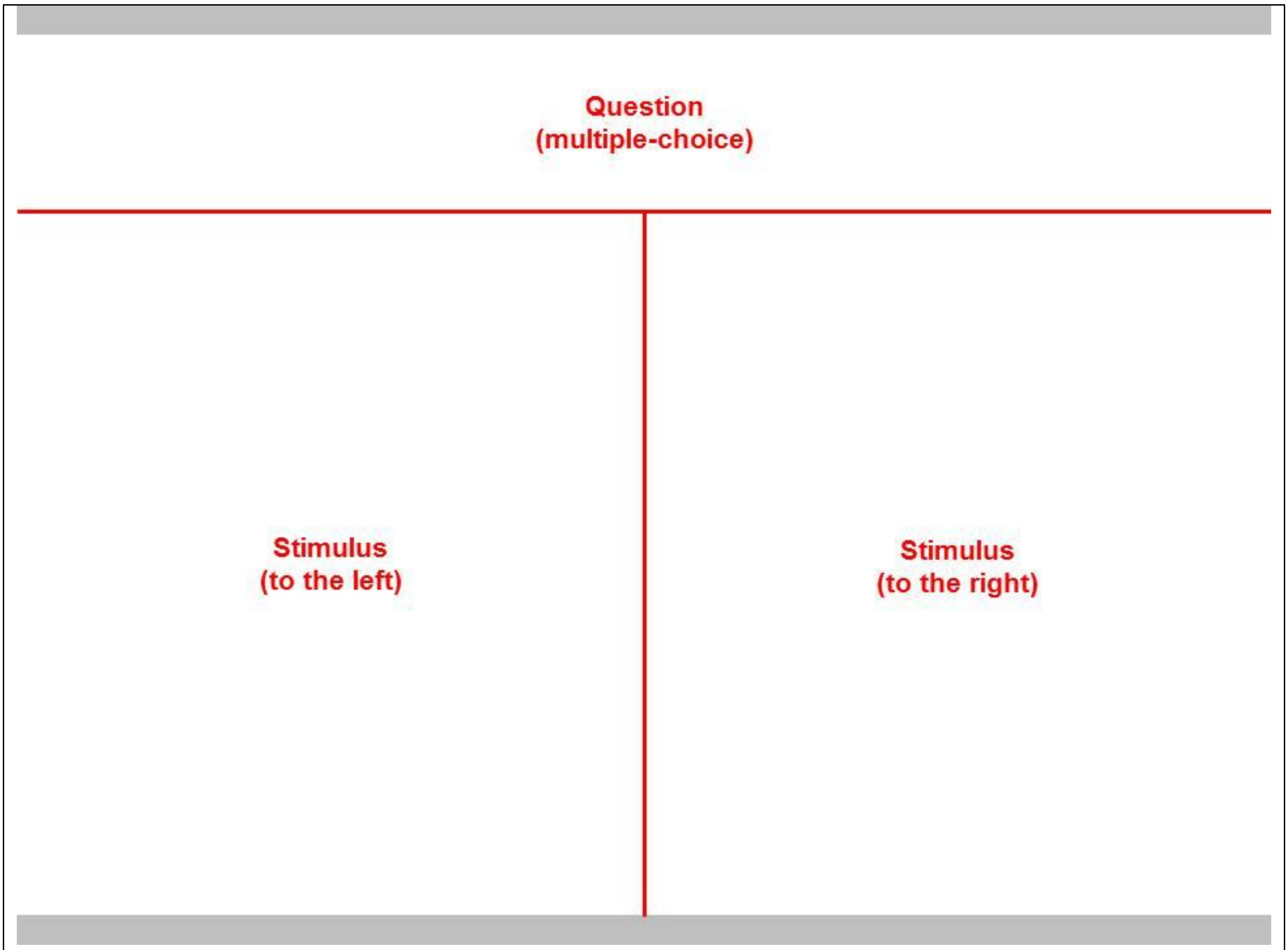


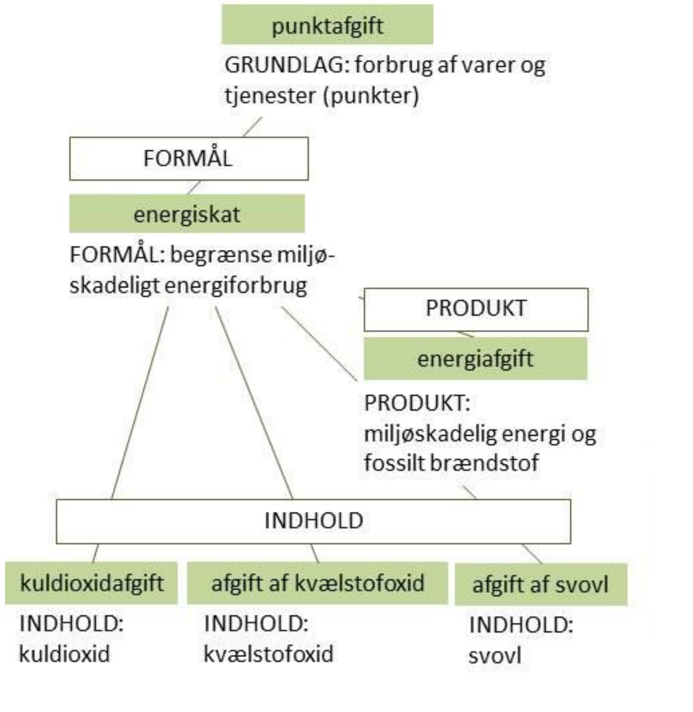
Figure D.2: Question-dual-entry-mode template. All questions were displayed in Danish in the experiment above the stimulus-pairs. The stimulus-pairs are separated by a vertical red line and the question is also separated from the stimulus-pairs by a horizontal red line. Six multiple-choice question types: The first diagram-based question (**D1**) is about subordinates and the second diagram-based question (**D2**) is about subdivision criteria. The first article-based question (**A1**) is about equivalence and the second article-based question (**A2**) is about comments. The first diagram-and-article-based question (**DA1**) is about type and the second diagram-and-article-based question (**DA2**) is about attributes.

The question-stimuli template is randomized at three levels: Firstly, the display-side of the diagram of the first question type (D1, A1 or DA1) is randomly assigned (right or left) assigning the opposite display-side to the second question type (D2, A2 or DA2). Secondly, the question types in each block are randomized. Finally, the blocks (B1 to B8) are randomly ordered throughout the experiment. The grey bars in the top and bottom were added to avoid distortion of the image when displayed on the computer connected to the eye-tracker. For an instance of a question and stimulus-pair, see Figure D.11.

Table D.1: Question types, questions and available answers of block 1. Translation into English:

D1: ”How many types of energy taxes exist? 1: Four; 2: Six; 3: Eight.”; **D2:** ”What separates carbondioxide tax from duty on nitrogen oxides? 1: Purpose; 2: Content; 3: Taxpayer.”; **A1:** ”What can ’energy tax’ be translated into in Danish? 1 (e*nergiafgift*): ’energy duty’ ; 2 (*energiskat*): ’energy tax’; 3 (*energitakst*): ’energy rate’.”; **A2:** ”Energy taxes constituted 44 per cent of excise duties in 2011 according to whom? 1: OECD; 2: Eurostat; 3: Statistics Denmark.”; **DA1:** ”What type of tax or duty is energy tax? 1: Environmental duty; 2: Energy duty; 3: Excise duty.”; **DA2:** ”What is the purpose of energy tax? 1: Limiting environmentally damaging energy consumption; 2: Limiting environmentally damaging consumption; 3: Limiting environmental damage.”

D1: Hvor mange typer energiskatter findes der? 1. Fire; 2. Seks; 3. Otte.
D2: Hvad adskiller kuldioxidafgift fra afgift af kvælstofoxid? 1. Formål; 2. Indhold; 3. Skatteyder.
A1: Hvad kan ‘energy tax’ oversættes til på dansk? 1. Energiafgift; 2. Energiskat; 3. Energitakst.
A2: I 2011 udgjorde energiskatterne 44 pct. af punktafgifterne ifølge hvem? 1. OECD; 2. Eurostat; 3. Danmarks Statistik.
DA1: Hvilken type skat eller afgift er energiskat? 1. Miljøafgift; 2. Energiafgift; 3. Punktafgift.
DA2: Hvad er formålet med energiskat? 1. Begrænse miljøskadeligt energiforbrug; 2. Begrænse miljøskadeligt forbrug; 3. Begrænse miljøbelastningen.



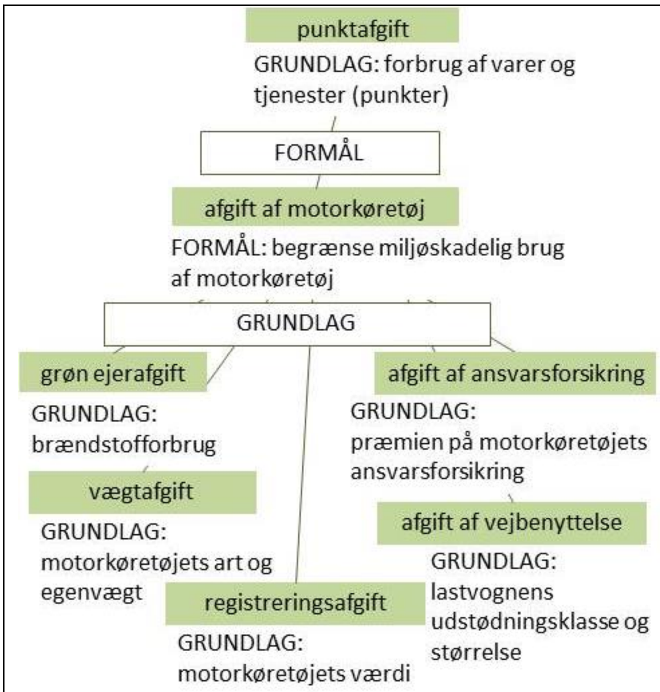
Dansk:	**energiskat**
Kilde til term:	Skatter og afgifter 2012 (Danmarks Statistik)
Definition:	punktafgift som pålægges varer og tjenester med det formål at begrænse miljøskadeligt energiforbrug
Kilde:	DanTermBank
Kommentar:	Den største gruppe blandt punkt-afgifterne er energiskatterne, der i 2011 tegnede sig for 44 pct. af samtlige punktafgifter.
Kilde:	Skatter og afgifter 2012 (Danmarks Statistik)
English:	**energy tax**
Source for term:	Taxation trends in the EU 2012 (Eurostat)
Definition:	excise duty imposed on energy consumption to reduce emissions
Source:	DanTermBank

Figure D.3: Block-1 dual-entry mode for *'energiskat'*. To the left: Concept diagram of block 1. To the right: Article entry of block 1. Translation into English:

Diagram: Superordinate (*punktafgift*): 'excise duty' with subdivision criteria: TAX BASE: consumption of goods and services. **Entry** (*energiskat*): 'energy tax' with subdivision criteria: PURPOSE: limiting environmentally damaging energy consumption. **Subordinate** (*energiafgift*): 'energy duty' with subdivision criteria: PRODUCT: environmentally damaging energy or fossile fuels. **Subordinate** (*kuldioxidafgift*): 'carbon dioxide tax' with subdivision criteria: CONTENT: carbon dioxide. **Subordinate** (*afgift af kvælstofoxid*): 'duty on nitrogen oxides' with subdivision criteria: CONTENT: nitrogen oxides. **Subordinate** (*afgift af svovl*): 'duty on sulphur' with subdivision criteria: CONTENT: sulphur.

Article: Row 1: Danish: energy tax. **Row 2:** Source for term: Taxes and Duties 2012 (Statistics Denmark). **Row 3:** Definition: excise duty on goods and services with the purpose of limiting the environmentally damaging energy consumption. **Row 4:** Source: DanTermBank. **Row 5:** Comment: The largest subgroup of the excise duties is energy taxes, which constituted 44 per cent of total excise duties in 2011. **Row 6:** Source: Taxes and Duties 2012 (Statistics Denmark).

Table D.2: Question types, questions and available answers of block 2. Translation into English: **D1:** ”How many types of motor vehicles taxes exist? 1: Three; 2: Four; 3: Five.”; **D2:** ”What separates grøn ejerafgift from vægtafgift? 1: Tax base; 2: Taxpayer; 3: ’Green tax is synonymous with ’weight tax’.”; **A1:** ”What can ’motor vehicles tax’ be translated into in Danish? 1 (*afgift af motorkøretøj*): ’motor vehicles tax’; 2 (*registreringsafgift*): ’motor vehicle registration duty’; 3 (*trængselsafgift*): ’duty to prevent crowding’; **A2:** ”Where is the taxation of motor vehicles significant according to Eurostat? 1: Denmark, Sweden, Norway; 2: Denmark, Ireland, Cyprus, Malta; 3: Germany, The Netherlands, Belgium.”; **DA1:** ”What type of tax or duty is motor vehicles tax? 1: Environmental duty; 2: Excise duty; 3: Energy tax.”; **DA2:** ”What is the purpose of motor vehicles tax? 1: Limiting environmental damage; 2: Limiting environmentally damaging energy consumption; 3: Limiting environmentally damaging use of motor vehicles.”
D1: Hvor mange typer afgifter af motorkøretøj findes der? 1. Tre; 2. Fire; 3. Fem. D2: Hvad adskiller grøn ejerafgift fra vægtafgift? 1. Grundlag; 2. Skatteyder; 3. Grøn ejerafgift er synonym med vægtafgift. A1: Hvad kan ‘motor vehicles tax’ oversættes til på dansk? 1. Afgift af motorkøretøj; 2. Registreringsafgift; 3. Trængselsafgift. A2: Hvor er beskatningen af motorkøretøjer betydelig ifølge Eurostat? 1. Danmark, Sverige, Norge; 2. Danmark, Irland, Cypern, Malta; 3. Tyskland, Nederlandene, Belgien. DA1: Hvilken type skat eller afgift er afgift af motorkøretøj? 1. Miljøafgift; 2. Punktafgift; 3. Energiskat. DA2: Hvad er formålet med afgift af motorkøretøj? 1. Begrænse miljøbelastningen; 2. Begrænse miljøskadeligt energiforbrug; 3. Begrænse miljøskadelig brug af motorkøretøjer.



Dansk:	**afgift af motorkøretøj**
Kilde til term:	Den juridiske vejledning 2012-2 (SKAT)
Definition:	punktafgift som pålægges motorkøretøjer med det formål at begrænse miljøskadelig brug
Kilde:	DanTermBank
English:	**motor vehicles tax**
Source for term:	Taxation trends in the EU 2012 (Eurostat)
Definition:	tax imposed on transport vehicles
Source:	DanTermBank
Comment:	In some countries the contribution of taxes on vehicles is significant: for instance, in Denmark, Ireland, Cyprus and Malta they account for between 36 % and 44 % of environmental taxes.
Source:	Taxation trends in the EU 2012 (Eurostat)

Figure D.4: Block-2 dual-entry mode for '*afgift af motorkøretøj*'. To the left: Concept diagram of block 2. To the right: Article entry of block 2. Translation into English:

Diagram: Superordinate (*punktafgift*): 'excise duty' with subdivision criteria: TAX BASE: consumption of goods and services. **Entry** (*afgift af motorkøretøj*): 'motor vehicles tax' with subdivision criteria: PURPOSE: limiting environmentally damaging use of vehicles. **Subordinate** (*grøn ejerafgift*): 'green vehicle excise duty' with subdivision criteria: TAX BASE: fuel consumption. **Subordinate** (*vægtafgift*): 'weight tax' with subdivision criteria: TAX BASE: the type and weight of the vehicle. **Subordinate** (*registreringsafgift*): 'motor vehicle registration duty' with subdivision criteria: TAX BASE: the value of the vehicle. **Subordinate** (*afgift af vejbenyttelse*): 'road charges' with subdivision criteria: TAX BASE: truck size and exhaust-emission class. **Subordinate** (*afgift af ansvarsforsikring*): 'duty on third-party liability insurance' with subdivision criteria: TAX BASE: premium on third party liability insurance of the motor vehicle.

Article: Row 1: Danish: motor vehicles tax. **Row 2:** Source for term: The legal instruction 2012-2 (SKAT). **Row 3:** Definition: excise duty on vehicles with the purpose of limiting the environmentally damaging use. **Row 4:** Source: DanTermBank.

Table D.3: Question types, questions and available answers of block 3. Translation into English: **D1:** ”How many types of green taxes exist apart from energy duties? 1: One; 2: Two; 3: Three.”; **D2:** ”What separates energy duty from weight tax? 1: Energy duty is synonymous with environmental duty; 2: Purpose; 3: Tax base.”; **A1:** ”What can ’green tax’ be translated into in Danish? 1 (*grøn takst*): ’green charge; 2 (*grøn ejerafgift*): ’ ’green vehicle excise duty’; 3 (*grøn afgift*): ’green tax.”; **A2:** ”Which country has the highest level of green taxation in the EU according to Eurostat? 1: Denmark and The Netherlands; 2: Denmark and Sweden; 3: Denmark and Norway.”; **DA1:** ”What type of tax or duty is green tax? 1: Excise duty; 2: Energy duty; 3: Environmental duty.”; **DA2:** ”What is the purpose of green taxes? 1: Limiting environmentally damaging energy consumption; 2: Limiting environmentally damaging use and consumption; 3: Limiting environmentally damaging emission.”
D1: Hvor mange typer grønne afgifter findes der udover energiafgifterne? 1. En; 2. To; 3. Tre. D2: Hvad adskiller energiafgift fra miljøafgift? 1. Energiafgift er synonym med miljøafgift; 2. Formål; 3. Grundlag. A1: Hvad kan ‘green tax’ oversættes til på dansk? 1. Grøn takst; 2. Grøn ejerafgift; 3. Grøn afgift. A2: Hvem har ifølge Eurostat det højeste niveau af grøn beskatning i EU? 1. Danmark og Nederlandene; 2. Danmark og Sverige; 3. Danmark og Norge. DA1: Hvilken type afgift er grøn afgift? 1. Punktafgift; 2. Energiafgift; 3. Miljøafgift. DA2: Hvad er formålet med grønne afgifter? 1. Begrænse miljøskadeligt energiforbrug; 2. Begrænse miljøskadelig brug of forbrug; 3. Begrænse miljøskadelig udledning.

Dansk:	**grøn afgift**
Kilde til term:	Retsinformation (Skatteministeriet)
Definition:	punktafgift som pålægges varer og tjenester med det formål at begrænse miljøskadeligt brug og forbrug
Kilde:	DanTermBank
English:	**green tax**
Source for term:	Taxation trends in the EU 2012 (Eurostat)
Definition:	tax imposed on polluting industries and the use of polluting substances
Source:	DanTermBank
Comment:	At 4.0 % in 2010, Denmark and the Netherlands display the highest level of 'green' taxes followed by Slovenia (3.6 %).
Source:	Taxation trends in the EU 2012 (Eurostat)

Figure D.5: Block-3 dual-entry mode for 'grøn afgift'. To the left: Concept diagram of block 3. To the right: Article entry of block 3. Translation into English:

Diagram: Superordinate (*punktafgift*): 'excise duty' with subdivision criteria: TAX BASE: consumption of goods and services. **Entry** (*grøn afgift*): 'green tax' with subdivision criteria: PURPOSE: limiting environmentally damaging use and consumption. **Subordinate** (*energiafgift*): 'energy duty' with subdivision criteria: TAX BASE: environmentally damaging energy consumption. **Subordinate** (*afgift af motorkøretøj*): 'motor vehicles tax' with subdivision criteria: TAX BASE: environmentally damaging use of vehicle. **Subordinate** (*miljøafgift*): 'environmental tax' with subdivision criteria: TAX BASE: environmentally damaging substances or scarce natural resources.

Article: Row 1: Danish: green tax. **Row 2:** Source for term: The Secretariat for Legal Information (The Ministry of Taxation). **Row 3:** Definition: excise duty on goods and services with the purpose of limiting the environmentally damaging use and consumption. **Row 4:** Source: DanTermBank.

Table D.4: Question types, questions and available answers of block 4. Translation into English: **D1:** ”How many types of excise duties exist? 1: Five; 2: Seven; 3: Ten.”; **D2:** ”What separates duty on stimulants from duty on gaming? 1: Tax base; 2: Purpose; 3: Duty on stimulants is synonymous with duty on gaming.”; **A1:** ”What can ’excise duty’ be translated into in Danish? 1 (*forbrugsafgift*): ’consumption duty’; 2 (*punktafgift*): ’excise duty’; 3 (*forbrugsskat*): ’consumption tax’.”; **A2:** ”How are excise duties imposed according to Statistics Denmark? 1: Generally (All goods); 2: Progressively (goods with a certain value); 3: Specifically (Only certain goods or services).”; **DA1:** ”What type of tax or duty is excise duty? 1: Direct tax; 2: General tax; 3: Indirect tax.”; **DA2:** ”On what tax base are excise duties imposed? 1: Consumption of goods and services; 2: Consumption of nydelsesmidler; 3: Consumption of energy.”
D1: Hvor mange typer punktafgifter findes der? 1. Fem; 2. Syv; 3. Ti. D2: Hvad adskiller afgift af nydelsesmidler fra afgift på spil? 1. Grundlag; 2. Formål; 3. Afgift af nydelsesmidler er synonym med afgift på spil. A1: Hvad kan ‘excise duty’ oversættes til på dansk? 1. Forbrugsafgift; 2. Punktafgift; 3. Forbrugsskat. A2: Hvordan pålægges punktafgifter ifølge Danmarks Statistik? 1. Generelt (Samtlige varer); 2. Progressivt (varer med en særlig værdi); 3. Specielt (kun enkelte varer og tjenestegrupper). DA1: Hvilken type skat eller afgift er punktafgift? 1. Direkte skat; 2. Generel skat; 3. Indirekte skat. DA2: På hvilket grundlag pålægges punktafgifter? 1. Forbrug af varer og tjenester; 2. Forbrug af nydelsesmidler; 3. Forbrug af energi.

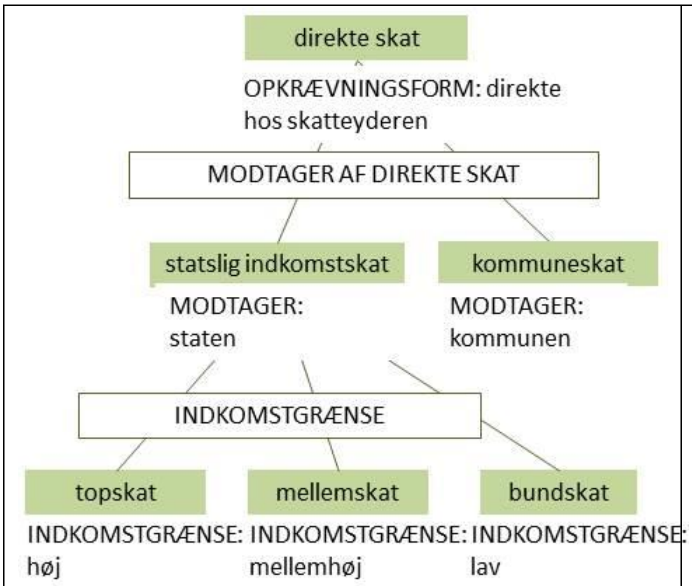
Dansk:	**punktafgift**
Kilde til term:	Den juridiske vejledning 2012-2 (SKAT)
Definition:	indirekte skat som pålægges forbrug af varer og tjenester med det formål at begrænse skadeligt forbrug
Kilde:	DanTermBank
Kommentar:	Afgifterne kan enten være pålagt generelt, dvs. omfatte samtlige varer og tjenester (moms), eller være pålagt specielt, dvs. kun vedrøre enkelte vare- eller tjenestegrupper (punktafgifter).
Kilde:	Skatter og afgifter 2012 (Danmarks Statistik)
English:	**excise duty**
Source for term:	Taxation trends in the EU 2012 (Eurostat)
Definition:	tax imposed on an act or transaction
Source:	DanTermBank

Figure D.6: Block-4 dual-entry mode for 'punktafgift'. To the left: Concept diagram of block 4. To the right: Article entry of block 4. Translation into English:

Diagram: Superordinate (*indirekte skat*): 'indirect tax' with subdivision criteria: COLLECTION FORM: indirectly on the transaction. **Entry** (*punktafgift*): 'excise duty' with subdivision criteria: TAX BASE: consumption of goods and services. **Subordinate** (*energiafgift*): 'energy duty' with subdivision criteria: PURPOSE: limiting environmentally damaging energy consumption. **Subordinate** (*afgift af motorkøretøj*): 'motor vehicles tax' with subdivision criteria: PURPOSE: limiting environmentally damaging vehicles. **Subordinate** (*miljøafgift*): 'environmental tax' with subdivision criteria: PURPOSE: limiting environmentally damaging products and contents. **Subordinate** (*afgift på spil*): 'duty on gaming' with subdivision criteria: PURPOSE: limiting consumption of games. **Subordinate** (*afgift af nydelsesmidler*): 'duty on stimulants' with subdivision criteria: PURPOSE: limiting unhealthy consumption of food, drink and tobacco.

Article: Row 1: Danish: excise duty. **Row 2:** Source for term: The legal instruction 2012-2 (SKAT). **Row 3:** Definition: indirect tax on the consumption of goods and services with the purpose of limiting damaging consumption. **Row 4:** Source: DanTermBank. **Row 5:** Comment: Duties are either imposed generally, i.e. including all goods and services (VAT), or they are imposed specifically, i.e. only including special groups of goods and services (excise duties). **Row 6:** Source: Taxes and Duties 2012 (Statistics Denmark).

Table D.5: Question types, questions and available answers of block 5. Translation into English: **D1:** ”How many types of state income taxes exist, if we include the middle-bracket tax? 1: One; 2: Two; 3: Three.”; **D2:** ”How is the middle-bracket tax collected? 1: Direct with the taxpayer (Direct tax); 2: Indirectly on the transaction (Indirect tax); 3: By the municipality (local income tax).”; **A1:** ”What can ’middle-bracket tax’ be translated into in Danish? 1 (*mellemskat*): ’middle-bracket tax’; 2 (*medieskat*): ’tax on media’; 3 (*middelskat*): ’mean tax’.”; **A2:** ”When was the middle-bracket tax abolished in Denmark according to Eurostat? 1: The tax reform of 2010; 2: The tax reform of 2009; 3: The middle-bracket tax has not been abolished.”; **DA1:** ”What type of tax is the middle-bracket tax? 1: Top-bracket tax; 2: Central government income tax; 3: Local income tax.”; **DA2:** ”What income limit is used to calculate the middle-bracket tax? 1: Low; 2: High; 3: Medium high.”
D1: Hvor mange typer statslige indkomstskatter findes der, hvis mellemskatten regnes med? 1. En; 2. To; 3. Tre. D2: Hvordan opkræves mellemskat? 1. Direkte hos skatteyderen (direkte skat); 2. Indirekte på transaktionen (indirekte skat); 3. Hos kommunen (kommuneskat). A1: Hvad kan ‘middle-bracket tax’ oversættes til på dansk? 1. Mellemskat; 2. Medieskat; 3. Middelskat. A2: Hvornår blev mellemskatten afskaffet i Danmark ifølge Eurostat? 1. Med skattereformen som blev vedaget i 2010; 2. Med skattereformen som blev vedtaget i 2009; 3. Mellemskatten er ikke blevet afskaffet. DA1: Hvilken type skat er mellemskat? 1. Topskat; 2. Statslig indkomstskat; 3. Kommuneskat. DA2: Hvilken indkomstgrænse anvendes ved beregning af mellemskatten? 1. Lav; 2. Høj; 3. Mellemhøj.



Dansk:	**mellemskat** *afskaffet*
Kilde til term:	Retsinformation (Skatteministeriet)
Definition:	statslig indkomstskat som udskrives på den del af indkomsten der over-stiger en mellemhøj indkomstgrænse
Kilde:	DanTermBank
English:	**middle-bracket tax**
Source for term:	Taxation trends in the EU 2012 (Eurostat)
Definition:	income tax imposed on the amounts exceeding the middle-bracket tax threshold
Source:	DanTermBank
Comment:	Since the 2009 tax cuts, which re-moved the middle tax bracket, the state income tax system has consisted of two tax brackets.
Source:	Taxation trends in the EU 2012 (Eurostat)

Figure D.7: Block-5 dual-entry mode for 'mellemskat'. To the left: Concept diagram of block 5. To the right: Article entry of block 5. Translation into English:

Diagram: Superordinate (*statslig indkomstskat*): 'state income tax' with subdivision criteria: RECEIVER: central government. **Entry** (*mellemskat*): 'middle-bracket tax' with subdivision criteria: INCOME LIMIT: medium. **Sideordinate** (*topskat*): 'top-bracket tax' with subdivision criteria: INCOME LIMIT: high. **Sideordinate** (*bundskat*): 'lower-bracket tax' with subdivision criteria: INCOME LIMIT: low. [**Super-sideordinate** (*kommuneskat*): 'local income tax' with subdivision criteria: RECEIVER: local government. **Super-superordinate** (*direkte skat*): 'direct tax' with subdivision criteria: COLLECTION FORM: directly with the taxpayer.]

Article: Row 1: Danish: middle-bracket tax. **Row 2:** Source for term: The Secretariat for Legal Information (The Ministry of Taxation). **Row 3:** Definition: state income tax imposed on the part of income above a medium-high income limit. **Row 4:** Source: DanTermBank.

Table D.6: Question types, questions and available answers of block 6. Translation into English: **D1:** ”How many types of real property taxes exist? 1: One; 2: Two; 3: Three.”; **D2:** ”What separates land value tax from reimbursement duty? 1: Land value tax is synonymous with reimbursement duty; 2: Revenue receiver; 3: Obligatory collection.”; **A1:** ”What can ’land tax’ be translated into in Danish? 1 (*grundskyld*): ’land value tax’; 2 (*ejendomsværdiskat*): ’property value tax’; 3 (*ejendomsskat*): ’land tax’.”; **A2:** ”From where does 84 per cent of the revenue of land taxation come from according to Statistics Denmark? 1: Land value tax; 2: Property value tax; 3: Reimbursement duty.”; **DA1:** ”What type of tax is land tax? 1: Direct tax; 2: Property value tax; 3: Reimbursement duty.”; **DA2:** ”On what tax base is land tax imposed? 1: Value of the property; 2: Value of the land; 3: Sum of the values of property and land.”
D1: Hvor mange typer ejendomsskatter findes der? 1. En; 2. To; 3. Tre. D2: Hvad adskiller grundskyld fra dækningsafgift? 1. Grundskyld er synonym med dækningsafgift; 2. Modtager af provenuet; 3. Obligatorisk at opkræve. A1: Hvad kan ‘land tax’ oversættes til på dansk? 1. Grundskyld; 2. Ejendomsværdiskat; 3. Ejendomsskat. A2: Hvorfra stammer 84 pct. af provenuet fra ejendomsbeskatningen ifølge Danmarks Statistik? 1. Grundskyld; 2. Ejendomsskyld; 3. Dækningsafgift. DA1: Hvilken type skat er ejendomsskat? 1. Direkte skat; 2. Ejendomsværdiskat; 3. Grundskyld. DA2: På hvilket grundlag pålægges ejendomsskatten? 1. Værdi af ejendom; 2. Værdi af grund; 3. Summen af værdien af ejendommen og grunden.

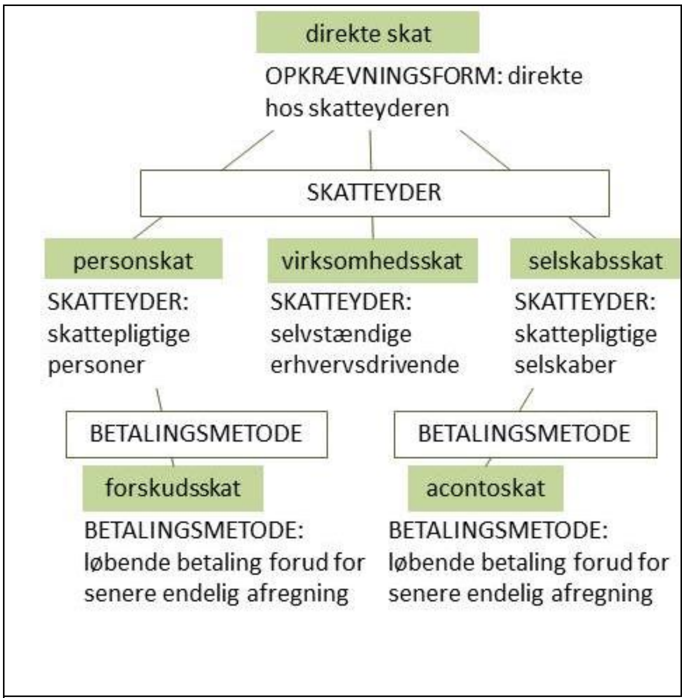
Dansk:	**ejendomsskat**
Kilde til term:	Retsinformation (Skatteministeriet)
Definition:	direkte skat som kommunen pålægger ejere af fast ejendom med udgangs-punkt i den offentlige vurdering af værdien af grunden
Kilde:	DanTermBank
Kommentar:	Grundskylden tegner sig i dag for over 84 pct. af ejendomsbeskatningen. Resten er dækningsafgifter.
Kilde:	Skatter og afgifter 2012 (Danmarks Statistik)
English:	**land tax**
Source for term:	Taxation trends in the EU 2012 (Eurostat)
Definition:	direct tax imposed on the assessed value of the land
Source:	DanTermBank

Figure D.8: Block-6 dual-entry mode for ’ejendomsskat’. To the left: Concept diagram of block 6. To the right: Article entry of block 6. Translation into English:

Diagram: Superordinate (*direkte skat*): ’direct tax’ with subdivision criteria: COLLECTION FORM: direct with the taxpayer. **Entry** (*ejendomsskat*): ’land tax’ with subdivision criteria: TAX BASE: land value. **Sideordinate** (*ejendomsværdiskat*): ’property value tax’ with subdivision criteria: TAX BASE: property value assessment. **Subordinate** (*grundskyld*): ’land value tax’ with subdivision criteria: OBLIGATORY: yes. **Subordinate** (*dækningsafgift*): ’reimbursement duty’ with subdivision criteria: OBLIGATORY: no.

Article: Row 1: Danish: land tax. **Row 2:** Source for term: Retsinformation (The Ministry of Taxation). **Row 3:** Definition: direct tax imposed on property owners by the local government based on the public property value assessment. **Row 4:** Source: DanTermBank. **Row 5:** Comment: Land value tax constitutes more than 84 per cent of the property taxation. The rest is reimbursement duty. **Row 6:** Source: Taxes and Duties 2012 (Statistics Denmark).

Table D.7: Question types, questions and available answers of block 7. Translation into English: **D1:** ”How many types of taxpayers exist apart from persons and companies? 1: One; 2: Two; 3: Three.”; **D2:** ”How do you pay personal taxes? 1: Annually; 2: Continual, preliminary; 3: Continual, backwards.”; **A1:** ”What can ’personal income tax’ be translated into in Danish? 1 (*personlig skat*): ’personal tax’; 2 (*personskat*): ’personal income tax’; 3 (*indkomstskat*): ’income tax’.”; **A2:** ”According to which law is the personal tax calculated, if we ask Statistics Denmark? 1: 'Tax Rate Act'; 2: 'Tax at Source Act'; 3: 'Personal Tax Act'”; **DA1:** ”What type of tax is personal tax? 1: Income tax; 2: Prepaid tax; 3: Direct tax.”; **DA2:** ”What type of taxpayers pay personal tax? 1: Liable persons; 2: Liable companies; 3: Self-employed.”
D1: Hvor mange typer skatteydere findes der udover personer og selskaber? 1. En; 2. To; 3. Tre. D2: Hvordan betales personskat? 1. Årligt; 2. Løbende, forud; 3. Løbende, bagud. A1: Hvad kan ‘personal income tax’ oversættes til på dansk? 1. Personlig skat; 2. Personskat; 3. Indkomstskat. A2: Efter hvilken lov opgøres personskatten ifølge Danmarks Statistik? 1. Udskrivningsprocentloven; 2. Kildeskatteloven; 3. Personskatteloven. DA1: Hvilken type skat er personskat? 1. Indkomstskat; 2. Acontoskat; 3. Direkte skat. DA2: Hvilken type skatteydere betaler personskat? 1. Skattepligtige personer; 2. Skattepligtige selskaber; 3. Selvstændige erhvervsdrivende.



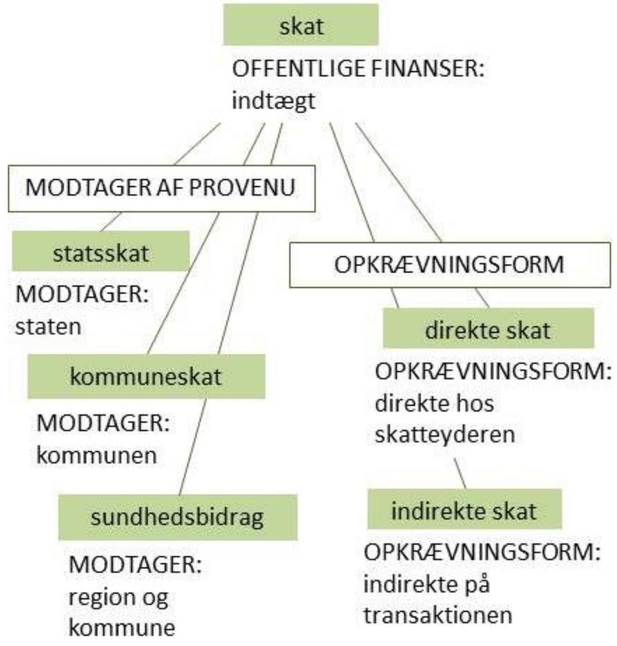
Dansk:	**personskat**
Kilde til term:	Retsinformation (Skatteministeriet)
Definition:	direkte skat som en skattepligtig person betaler på grundlag af sin skattepligtige indkomst eller formue DanTermBank
Kilde:	
Kommentar:	I personskatteloven fastsættes reglerne for opgørelse af personlig indkomst, kapitalindkomst samt de fradrag, der kan foretages heri.
Kilde:	Skatter og afgifter 2012 (Danmarks Statistik)
English:	**personal income tax**
Source for term:	Taxation trends in the EU 2012 (Eurostat)
Definition:	direct tax imposed on income
Source:	DanTermBank

Figure D.9: Block-7 dual-entry mode for 'personskat'. To the left: Concept diagram of block 7. To the right: Article entry of block 7. Translation into English:

Diagram: Superordinate (*direkte skat*): 'direct tax' with subdivision criteria: COLLECTION FORM: directly with the taxpayer. **Entry** (*personskat*): 'personal income tax' with subdivision criteria: TAXPAYER: taxable persons. **Sideordinate** (*virksomhedsskat*): 'corporation tax' with subdivision criteria: TAXPAYER: self-employed persons. **Sideordinate** (*selskabsskat*): 'corporate tax' with subdivision criteria: TAXPAYER: taxable companies. **Subordinate** (*forskudsskat*): 'provisional tax' with subdivision criteria: PAYMENT: continuous, in advance of final settlement. **Subordinate** (*acontoskat*): 'prepaid tax' with subdivision criteria: PAYMENT: continuous, in advance of final settlement.

Article: Row 1: Danish: personal income tax. **Row 2:** Source for term: The Secretariat for Legal Information (The Ministry of Taxation). **Row 3:** Definition: direct tax payed by taxable persons based on the taxable income or wealth. **Row 4:** Source: DanTermBank. **Row 5:** Comment: In the Personal tax law the rules for assessing the personal income, capital income as well as the deductions. **Row 6:** Source: Taxes and Duties 2012 (Statistics Denmark).

Table D.8: Question types, questions and available answers of block 8. Translation into English: **D1:** ”What types of collection forms exist? 1: None; 2: Only a direct form; 3: A direct and an indirect form.”; **D2:** ”How do direct taxes contribute to the public finances? 1: As income; 2: As expenditure; 3: Direct taxes do not constitute a part of the public finances.”; **A1:** ”What can ’direct tax’ be translated into in Danish? 1 (*direkte skat*): ’direct tax’; 2 (*direkte afgift*): ’direct duty’; 3 (*indkomstskat*): ’income tax’.”; **A2:** ”What taxes are best suited for redistribution according to Eurostat? 1: Indirect taxes; 2: Direct taxes; 3: It is not possible to use taxes for redistribution purposes.”; **DA1:** ”What type of tax, VAT or duty is direct tax? 1: VAT; 2: Tax; 3: Duty.”; **DA2:** ”How are direct taxes imposed? 1: On the transaction; 2: On the value increase; 3: With the taxpayer.”
D1: Hvilke typer opkrævningsformer findes der? 1. Ingen; 2. Kun en direkte form; 3. Både en direkte og en indirekte form. D2: Hvordan bidrager de direkte skatter til de offentlige finanser? 1. Som en indtægt; 2. Som en udgift; 3. Direkte skatter er ikke en del af de offentlige finanser. A1: Hvad kan ‘direct tax’ oversættes til på dansk? 1. Direkte skat; 2. Direkte afgift; 3. Indkomstskat. A2: Hvilke skatter egner sig bedst til omfordeling ifølge Eurostat? 1. Indirekte skatter; 2. Direkte skatter; 3. Skatter kan ikke benyttes til omfordeling. DA1: Hvilken type skat, moms eller afgift er direkte skat? 1. Moms; 2. Skat; 3. Afgift. DA2: Hvordan opkræves direkte skatter? 1. På transaktionen; 2. På værdiforøgelsen; 3. Hos skatteyderen.



Dansk:	**direkte skat**
Kilde til term:	Fordeling og incitamenter 2004 (Finansministeriet)
Definition:	skat som det offentlige pålægger skatteyderens indkomst eller formue
Kilde:	DanTermBank
English:	**direct tax**
Source for term:	Taxation trends in the EU 2012 (Eurostat)
Definition:	tax imposed on the taxpayer's personal income and net worth
Source:	DanTermBank
Comment:	Taxes are traditionally classified as direct or indirect; generally, the first group allows greater redistribution as it is impractical to introduce progressivity in indirect taxes.
Source:	Taxation trends in the EU 2012 (Eurostat)

Figure D.10: Block-8 dual-entry mode for 'direkte skat'. To the left: Concept diagram of block 8. To the right: Article entry of block 8. Translation into English:

Diagram: Superordinate (*skat*): 'tax' with subdivision criteria: PUBLIC FINANCES: income. **Entry** (*direkte skat*): 'direct tax' with subdivision criteria: COLLECTION FORM: direct with the taxpayer. **Sideordinate** (*indirekte skat*): 'indirect tax' with subdivision criteria: COLLECTION FORM: indirectly on the transaction. **Sideordinate** (*statsskat*): 'central government tax' with subdivision criteria: RECEIVER: central government. **Sideordinate** (*kommuneskat*): 'local state tax' with subdivision criteria: RECEIVER: local government. **Sideordinate** (*sundhedsbidrag*): 'healthcare contribution' with subdivision criteria: RECEIVER: local authorities.

Article: Row 1: Danish: direct tax. **Row 2:** Source for term: "Distribution and Incentives 2004" (The Ministry of Finance). **Row 3:** Definition: tax imposed taxpayers' income or wealth. **Row 4:** Source: DanTermBank.

Appendix E:
Retrospective interview

Table E.1: Retrospective interview. Questions (1 to 15) in the retrospective interview were answered by participants upon the completion of the eye-tracking experiment. Translation into English:

Section 1: ”You have completed the experiment, and I would like to close the experiment down with a couple of questions. I would like you to use a seven-point scale from 1-7. I would also like you to qualify your answers.”

Section 2: ”User situation. **1:** On a scale from 1-7, where 1 is the poorest and 7 is the best, how well do you think the experiment resembles a natural user situation (i.e. could you imagine a real situation, where you would have similiar questions)? **2:** And how well do you think you got your information need covered (i.e. did the articles and/or diagrams provide you with an answer to the posed questions? Please, qualify.”

Section 3: ”Articles and diagrams cover different information needs. Imagine that you are looking for a definition of a technical term.” **3:** ”On a scale from 1-7, where 1 is the poorest and 7 is the best, how effective do you think it is to seek information (i.e. find a definition) schematically in an article?” **4:** ”How effective do you think it is to seek information (i.e. find a definition) graphically in a diagram?” **5:** ”How well do you think that you performed when you retrieved answers in articles?” **6:** ”How well do you think that you performed when you retrieved answers in diagrams? Please, qualify.”

Section 4: ”Specialized language and technical terms.” **7:** ”On a scale from 1-7, where 1 is the most easy and 7 is the most difficult, how difficult do you think the specialized text from Statistics Denmark was? If you think it was difficult, was it primarily due to the terms? Please, qualify.” **8:** ”Which technical term(s) from the text, do you find the most and least difficult? Please, qualify.” **9:** ”Which existing technical term(s) from the list, do you find the most and least difficult? Please, qualify.” **10:** ”On a scale from 1-7, where 1 is never and 7 is often, how often do you read or write Danish specialized texts covering the taxation field in your current job? Please, qualify” **11:** ”On a scale from 1-7, where 1 is never and 7 is often, how often do you read or write English specialized texts covering the taxation field in your current job? Please, qualify.”

Section 5: ”Imagine a publicly accessible term bank, i.e. a database containing structured information about technical terms from many different subject fields (not only the taxation field). **12:** ”On a scale from 1-7, where 1 is not relevant and 7 is very relevant, how relevant would a termbank allowing you to search specialized knowledge by means of articles and/or diagrams be for your daily work?” **13:** ”Would you prefer to start your search where information is presented in an article, as a diagram, or do you not have special preferences? Please, qualify.” **14:** ”On a scale from 1-7, where 1 is never and 7 is often, how often do you need to conduct concept clarification in Danish in your current job? Please, qualify.” **15:** ” On a scale from 1-7, where 1 is never and 7 is often, how often do you need to conduct concept clarification in English in your current job? Please, qualify.”

Section 6: ”Thank you for participating!”

Du har nu gennemført forsøget, og jeg vil gerne slutte af med at stille dig et par spørgsmål. Jeg vil bede dig bruge en **skala** fra 1-7. Jeg vil også meget gerne have en **begrundelse** for dine svar.
Brugersituation 1. På en skala fra 1-7, hvor 1 er dårligst og 7 er bedst, hvor godt synes du, at forsøget afspejler en naturlig brugersituation (dvs. kunne du forestille dig en virkelig situation, hvor du ville stå med lignende spørgsmål)? 2. Og hvor godt synes du, at du fik du dækket dit informationsbehov (dvs. gav artiklerne og/eller diagrammerne svar på de stillede spørgsmål)? Begrund:____ ____
Artikler og diagrammer dækker forskellige informationsbehov. Forestil dig nu, at du skal finde en definition på et fagudtryk. 3. På en skala fra 1-7, hvor 1 er dårligst og 7 er bedst, hvor godt synes du, det er at søge information (dvs. finde en definition), som er opstillet skematisk form i en **artikel**? 4. Hvor godt synes du, det er at søge information (dvs. finde en definition), som er opstillet på grafisk form i et **diagram**? 5. Og hvor godt synes du, at du klarede opgaverne, hvor du fandt svaret i en **artikel**? 6. Og hvor godt synes du, at du klarede opgaverne, hvor du fandt svaret i et **diagram**? Begrund:____ ____

Fagsprog og fagudtryk

7. På en skala fra 1-7, hvor 1 er nemmest og 7 er sværest, hvor **svær** synes du, fagteksten fra Danmarks Statistik var?

- Hvis du synes, at fagteksten var svær: Var det så primært fagudtrykkene, som gjorde teksten svær?

Begrund:____

____

8. Hvilke fagudtryk fra **teksten**, synes du, er henholdsvis sværest og lettest:

Sværest:____

Lettest:____

Begrund:____

9. Hvilke eksisterende fagudtryk fra **listen**, synes du, er henholdsvis sværest og lettest:

Sværest:____

Lettest:____

Begrund:____

10. På en skala fra 1-7, hvor 1 er aldrig og 7 er ofte, hvor tit læser eller skriver du **danske** fagtekster, som berører skatteområdet, i dit nuværende arbejde?

Begrund/Antal:____

11. På en skala fra 1-7, hvor 1 er aldrig og 7 er ofte, hvor tit læser eller skriver du **engelske** fagtekster, som berører skatteområdet, i dit nuværende arbejde (1-7)?

Begrund/Antal:____

Forestil dig en **offentligt tilgængelig termbank**, dvs. en database, som indeholder strukturerede informationer om fagudtryk fra mange forskellige fagområder (ikke kun skatteområdet).

12. På en skala fra 1-7, hvor 1 er ikke relevant og 7 er yderst relevant, hvor **relevant** ville en termbank, hvor du kan udsøge faglig viden i enten artikler og/eller diagrammer, være for dit daglige arbejde?
13. Ville du **foretrække** at begynde din søgning, hvor informationen er præsenteret i en artikel, som et diagram, eller har du ingen særlige præferencer?
    Begrund:
    ____
14. På en skala fra 1-7, hvor 1 er aldrig og 7 er ofte, hvor ofte har du brug for at foretage begrebsafklaring på **dansk** i dit nuværende arbejde?
    Begrund/Antal:____
15. Og hvor ofte har du tilsvarende brug for at foretage begrebsafklaring på **engelsk** i dit nuværende arbejde?
    Begrund/Antal:____
    ____

Tak for hjælpen!

Appendix F:
Card-sorting

Table F.1: Card-sorting exercise. Instruction to the card-sorting exercise appearing at the beginning of workshop prior to any exposure to concept diagrams. Translation into English:

Section 1: ”Working area. Do you primarily work with law, economics or politics on a daily basis? Yes; No.”
Section 2: ”Sorting. Below you are given 32 technical terms from the taxation field ordered alphabetically. Please, draw a diagram showing the relations between each term (use page 2). If you are not able to fit in all the terms into one diagram, you are allowed to use several diagrams.”

Arbejdsområde

Arbejder du primært med jura, økonomi eller politik til daglig?

- ☐ Ja
- ☐ Nej

Systematisering

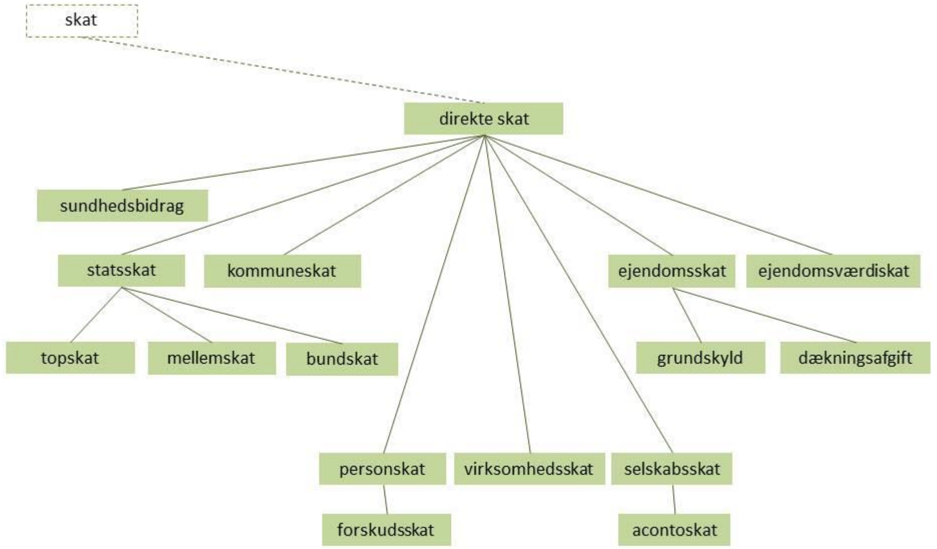
Nedenfor ser du 32 fagudtryk fra skatteområdet i alfabetisk rækkefølge.

Du skal nu tegne et diagram, som viser sammenhængen mellem de enkelte fagudtryk (benyt side 2).

Hvis du ikke kan få alle fagudtryk til at passe ind i ét diagram, må du gerne tegne flere.

acontoskat	grøn afgift
afgift af kvælstofoxid	grøn ejerafgift
afgift af motorkøretøj	indirekte skat
afgift af nydelsesmidler	kommuneskat
afgift af svovl	kuldioxidafgift
afgift af vejbenyttelse	mellemskat
afgift på spil	miljøafgift
bundskat	personskat
direkte skat	punktafgift
dækningsafgift	registreringsafgift
ejendomsskat	selskabsskat
ejendomsværdiskat	statsskat
energiafgift	sundhedsbidrag
energiskat	topskat
forskudsskat	virksomhedsskat
grundskyld	vægtafgift

Figure F.1: The given list of technical terms in the card-sorting exercise. The list comprises of 32 technical terms ordered alphabetically. The 32 terms were all the terms shown in the concept diagrams in the eye-tracking experiment (see left hand sides of figures D.3-D.10). Translation into English:
Term 1 (*acontoskat*): 'prepaid tax'; **Term 2** (*afgift af kvælstofoxid*): 'duty on nitrogen oxides'; **Term 3** (*afgift af motorkøretøj*): 'motor vehicle duties'; **Term 4** (*afgift af nydelsesmidler*): 'duty on stimulants'; **Term 5** (*afgift af svovl*): 'duty on sulpher'; **Term 6** (*afgift af vejbenyttelse*): 'road charges'; **Term 7** (*afgift på spil*): 'duty on gaming'; **Term 8** (*bundskat*): 'lower-bracket tax'; **Term 9** (*direkte skat*): 'direct tax'; **Term 10** (*dækningsafgift*): 'reimbursement duty'; **Term 11** (*ejendomsskat*): 'land tax'; **Term 12** (*ejendomsværdiskat*): 'property value tax'; **Term 13** (*energiafgift*): 'energy duty'; **Term 14** (*energiskat*): 'energy tax'; **Term 15** (*forskudsskat*): 'provisional tax'; **Term 16** (*grundskyld*): 'land value tax'; **Term 17** (*grøn afgift*): 'green tax'; **Term 18** (*grøn ejerafgift*): 'green vehicle excise duty'; **Term 19** (*indirekte skat*): 'indirect tax'; **Term 20** (*kommuneskat*): 'local state tax'; **Term 21** (*kuldioxidafgift*): 'carbon dioxide tax'; **Term 22** (*mellemskat*): 'middle-bracket tax'; **Term 23** (*miljøafgift*): 'environmental tax'; **Term 24** (*personskat*): 'personal income tax'; **Term 25** (*punktafgift*): 'excise duty'; **Term 26** (*registreringsafgift*): 'motor vehicle registration duty'; **Term 27** (*selskabsskat*): 'corporate tax'; **Term 28** (*statsskat*): 'central government tax'; **Term 29** (*sundhedsbidrag*): 'healthcare contribution'; **Term 30** (*topskat*): 'top-bracket tax'; **Term 31** (*virksomhedsskat*): 'corporation tax'; **Term 32** (*vægtafgift*): 'weight tax'.



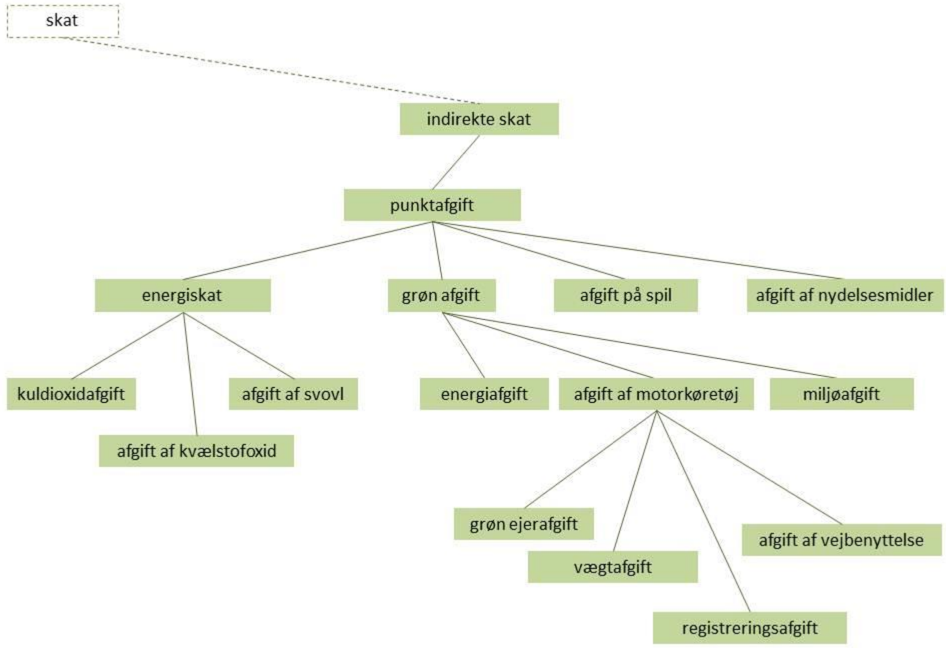


Figure F.2: Reference maps. Both maps contain 16 terms and 15 relations. In the upper part: Reference map of the direct taxes (*direkte skat*) from the given list of technical terms. In the lower part: Reference map of indirect taxes (*indirekte skat*) from the given list of technical terms. Dotted boxes contain terms that were not part of the given list of technical terms.

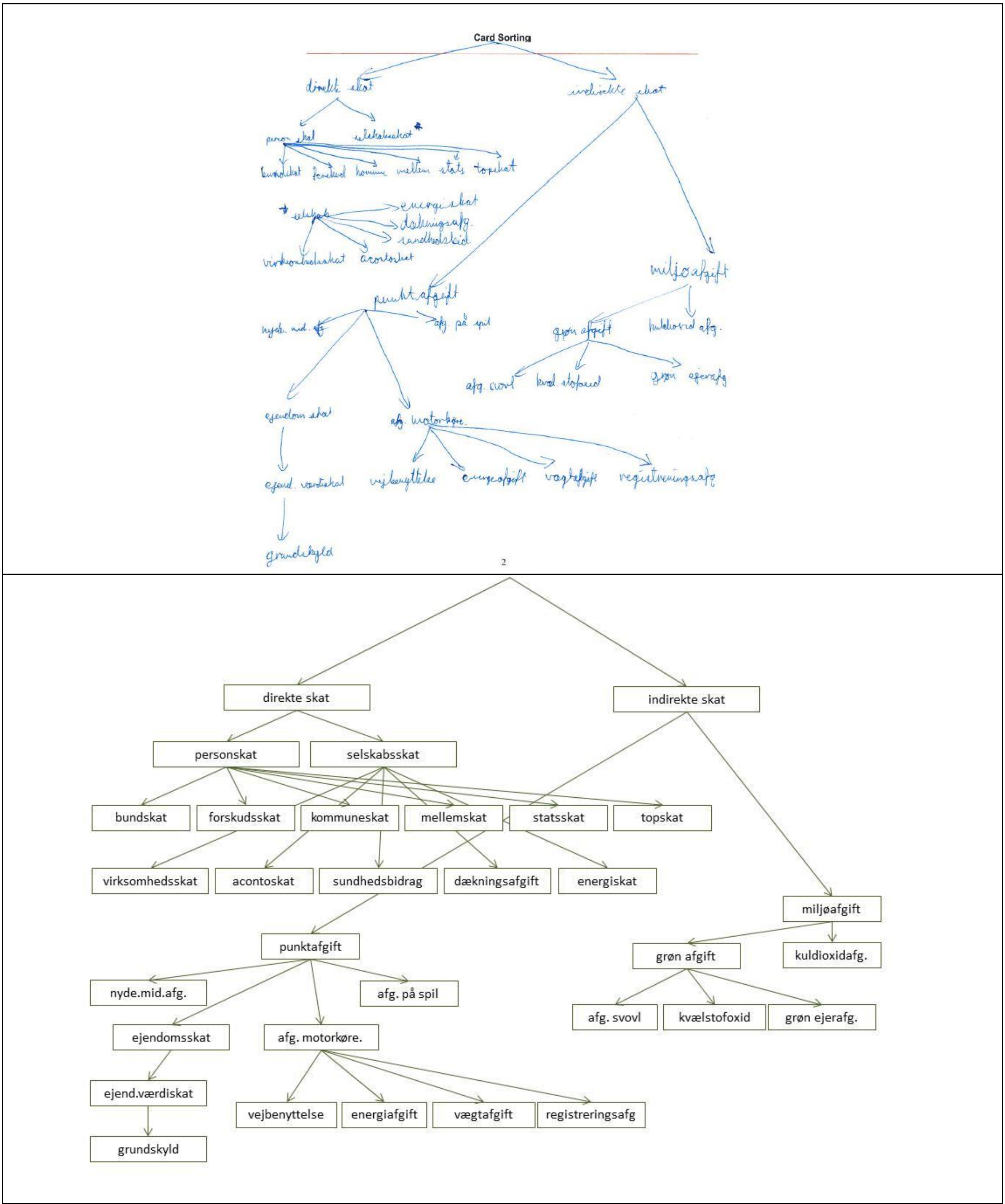


Figure F.3: In the upper part: Drawing made by participant 1. In the lower part: Reproduced drawing, which contains 1 diagram and 32 unique terms.

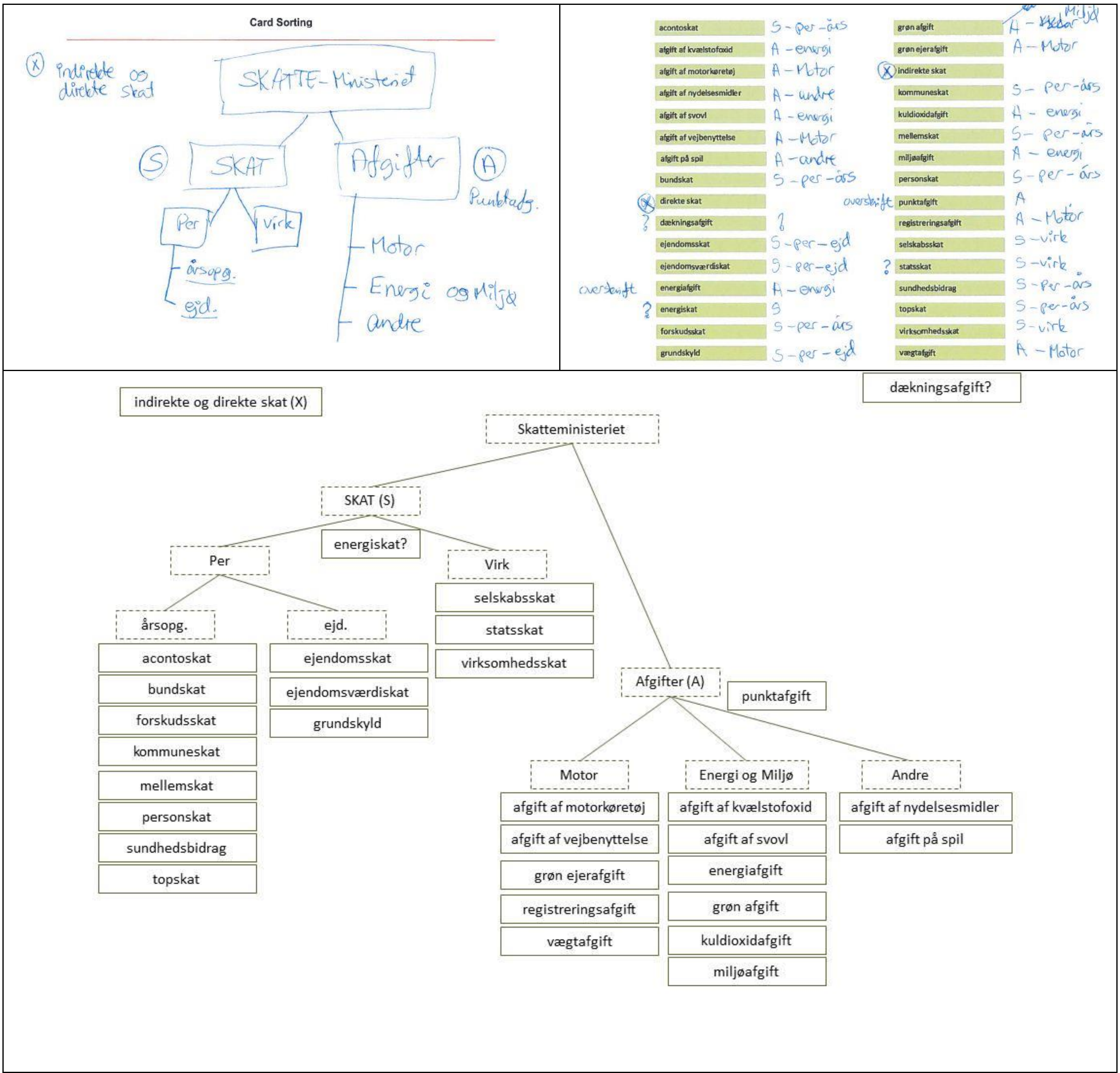


Figure F.4: In the upper part: Drawing made by participant 2. In the lower part: Reproduced drawing, which contains 1 diagram and 32 unique terms (including 2 terms placed outside the diagram). Dotted boxes contain terms that were not part of the given list of technical terms.

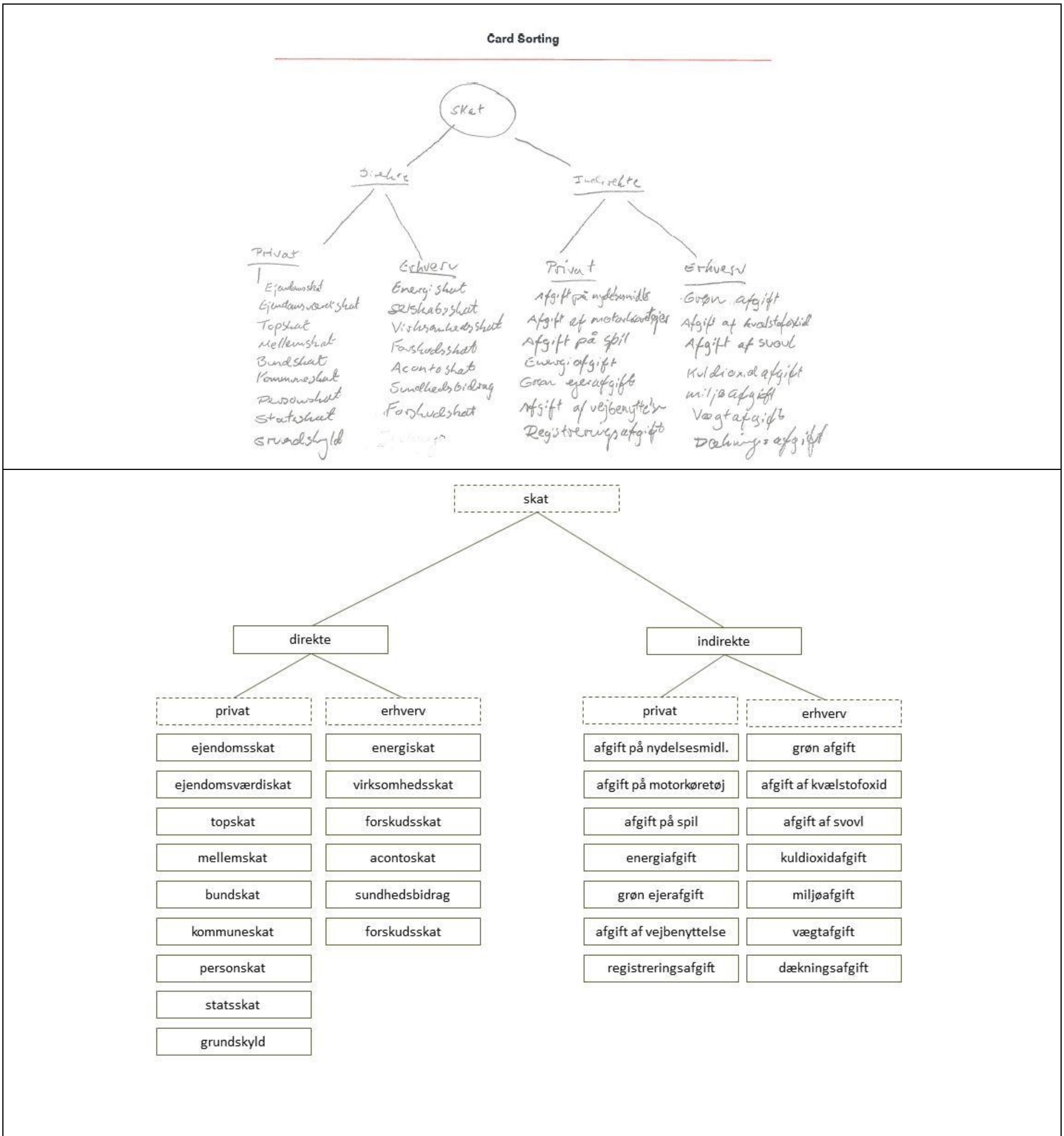


Figure F.5: In the upper part: Drawing made by participant 3. In the lower part: Reproduced drawing, which contains 1 diagram and 30 unique terms. Dotted boxes contain terms that were not part of the given list of technical terms.

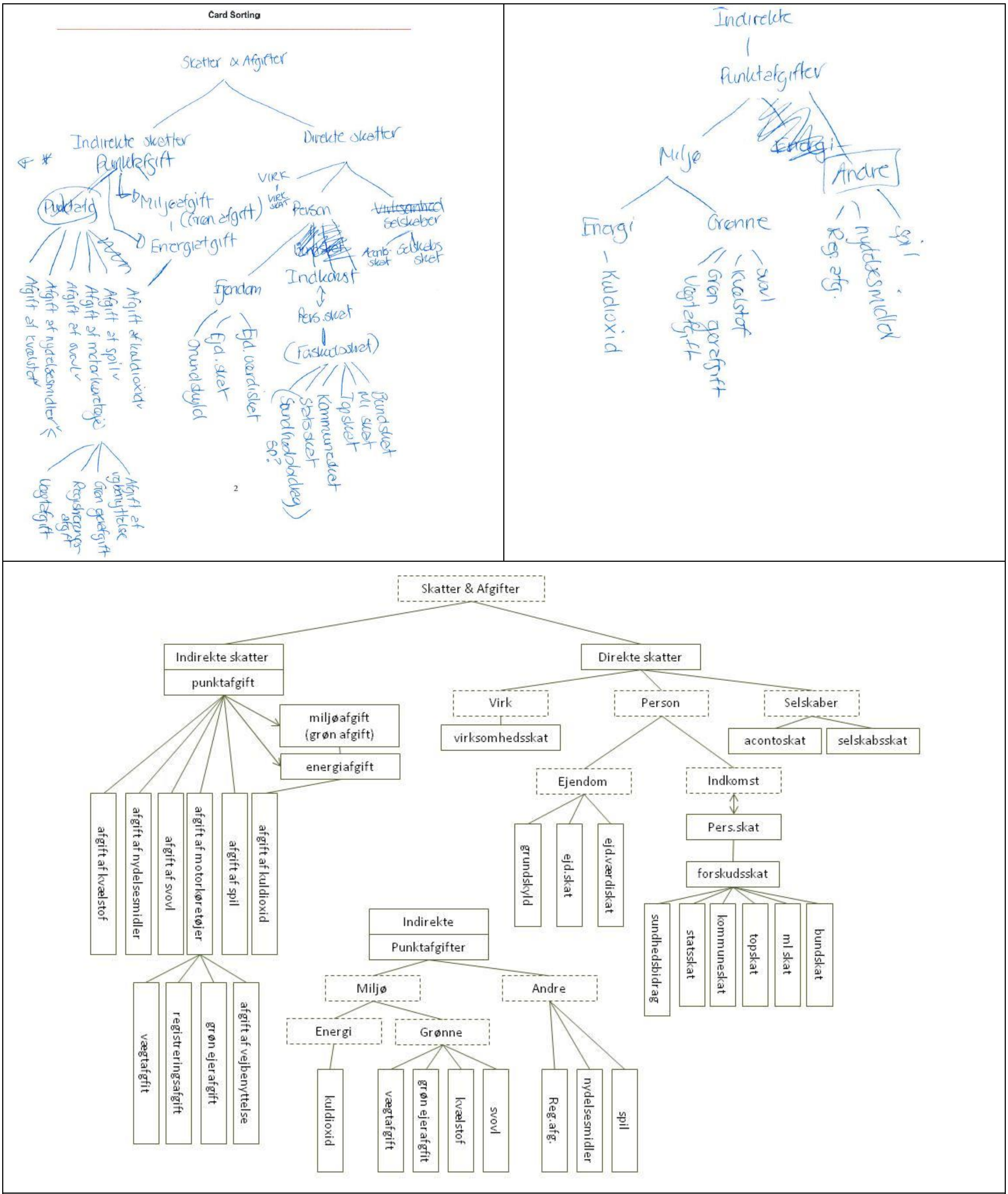


Figure F.6: In the upper part: Drawing made by participant 4. In the lower part: Reproduced drawing, which contains 2 diagrams and 30 unique terms. Dotted boxes contain terms that were not part of the given list of technical terms.

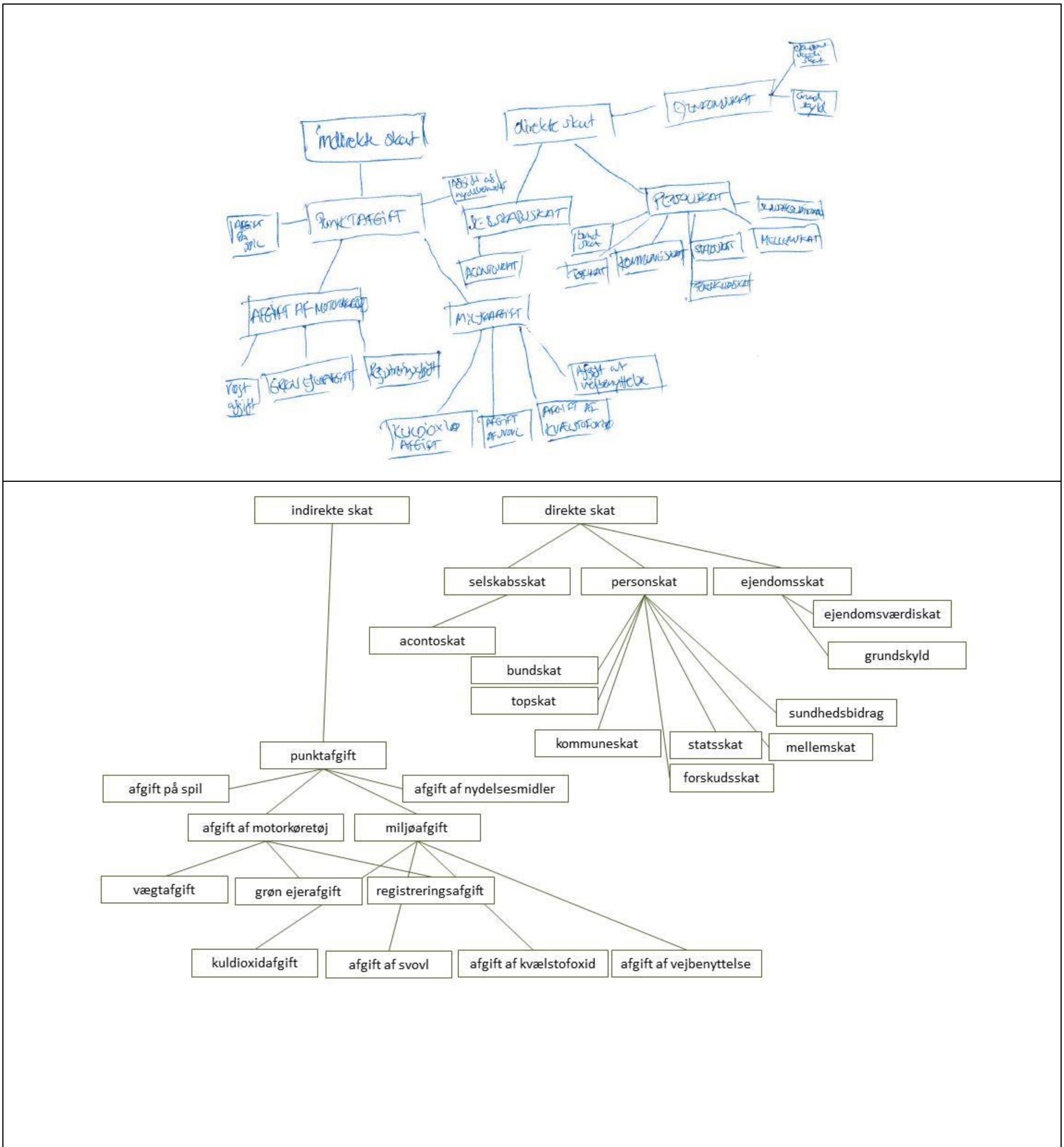
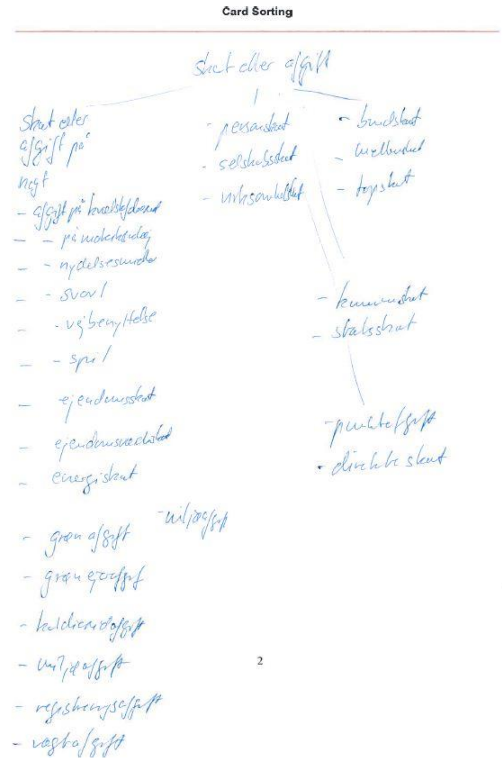


Figure F.7: In the upper part: Drawing made by participant 5. In the lower part: Reproduced drawing, which contains 2 diagrams and 27 unique terms.



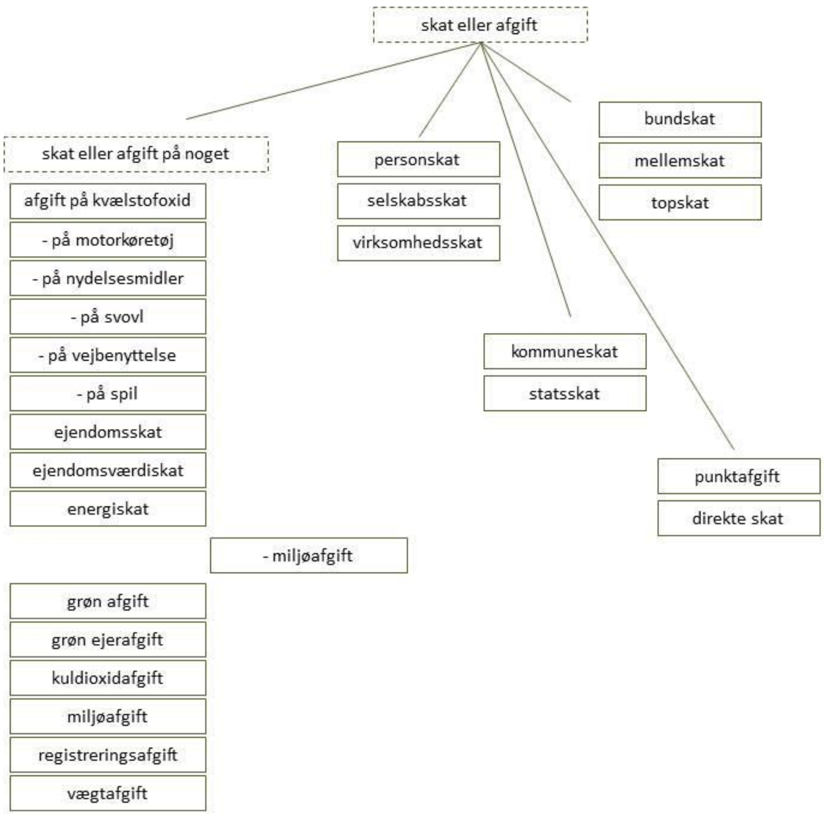


Figure F.8: In the upper part: Drawing made by participant 6. In the lower part: Reproduced drawing, which contains 1 diagram and 25 unique terms. Dotted boxes contain terms that were not part of the given list of technical terms.

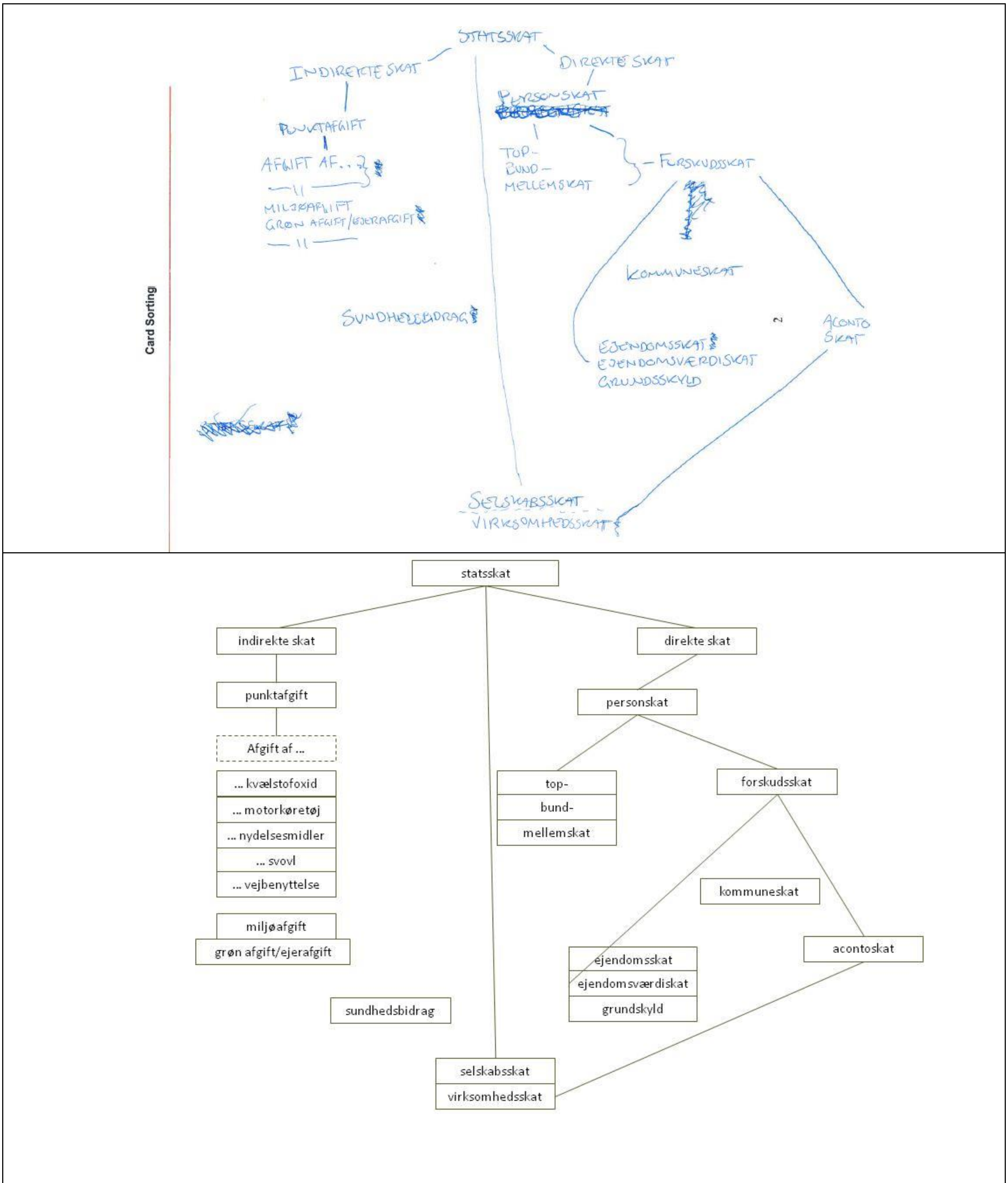
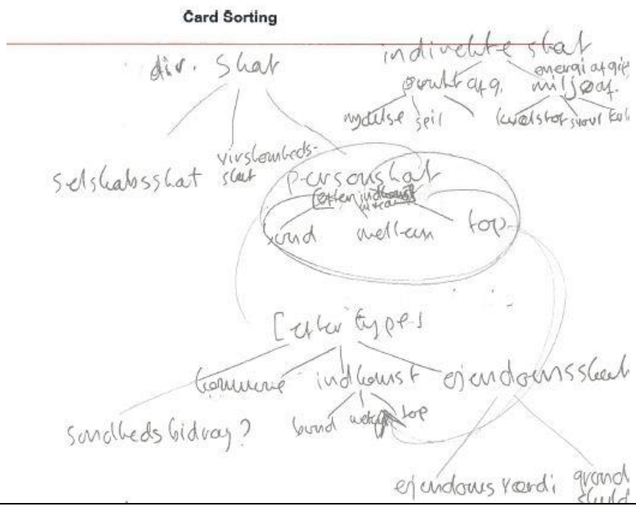


Figure F.9: In the upper part: Drawing made by participant 7. In the lower part: Reproduced drawing, which contains 2 diagrams and 25 unique terms. Dotted boxes contain terms that were not part of the given list of technical terms.



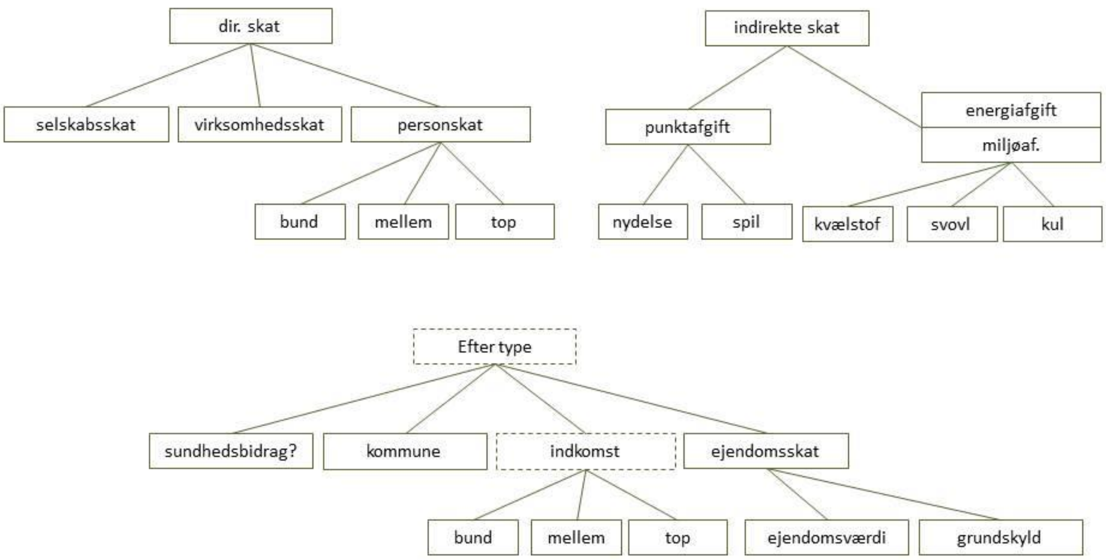


Figure F.10: In the upper part: Drawing made by participant 8. In the lower part: Reproduced drawing, which contains 3 diagrams and 21 unique terms. Dotted boxes contain terms that were not part of the given list of technical terms.

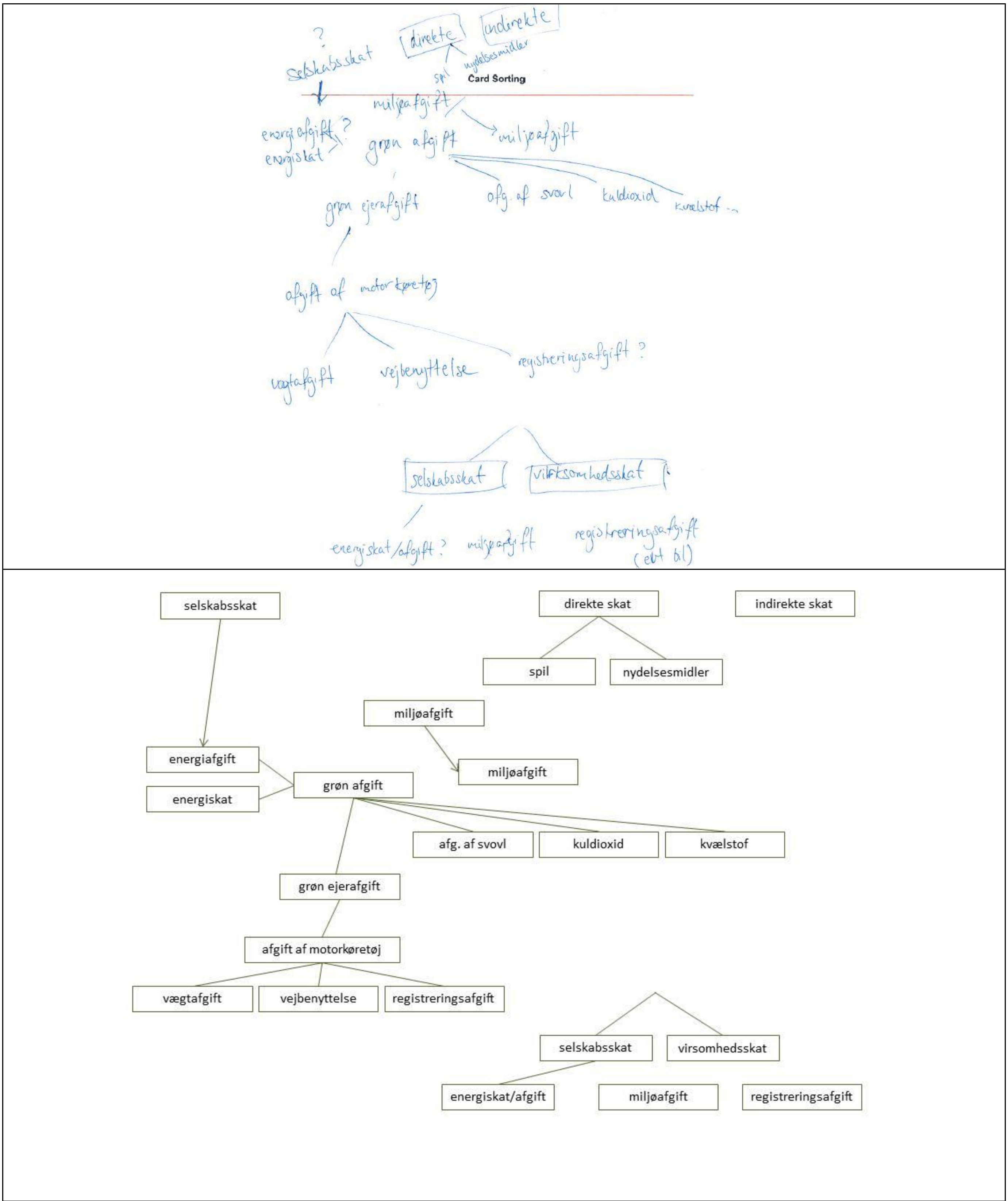


Figure F.11: In the upper part: Drawing made by participant 9. In the lower part: Reproduced drawing, which contains 4 diagrams and 18 unique terms (including 1 term placed outside the diagrams).

Appendix G:
Focus-group discussions

Table G.1: Instruction to discussion. Two slides used at the workshop showing the instruction to the focus-group discussions. Translation into English:

Section 1: "Individually, note on post-its: The 3 most important incentives that would make you use a termbase. The 3 most important barriers that would prevent you from using a termbase. The 3 most important effects on your work, if you had access to a termbase (Also please, indicate an effect measure)"

Section 2: "Group the post-its by identical content and evaluate the seriousness of incentives and barriers: Critical (most); Serious (medium); Less important (least)"

1. Indviduelt: Notere på post-its

De 3 vigtigste incitamenter for, at du vil tage en termbase i brug

De 3 vigtigste barrierer for, at du vil tage en termbase i brug

De 3 vigtigste effekter på dit arbejde, hvis du fik adgang til en termbase (angiv gerne en tilhørende effektmåling)

2. Grupper post-its efter identisk/beslægtet indhold + evaluer alvorsgraden af incitamenter og barrierer:

Kritisk (mest)

Alvorlig (mellem)

Mindre vigtig (mindst)

Table G.2: Focus-group conclusions. One slide per focus group was used to sum up (in plenum) conclusions from each of the four focus-group discussions. I = Incentives to use a termbase. B = Barriers against using a termbase. E = Effects to be gained from using a termbase. Translation into English:

Group 1: "The (user-friendly) web-editors.
I: Understand concepts without help from colleagues (Precision) + Ensuring the correct term
B: Knowing of its existence + Too seldom use + Recognize a need
E: Overview of the taxation law (visual) + Consensus on terms + Tool for writing texts"

Group 2: "The tax experts.
I: Ensure correct use of terms (critical) + Access to explanations (to laymen) + provide overview
B: Trust in up-to-date information coverage (and exclusions) + information overload
E: Time saving: Training + Consistent communication + Common understanding across disciplines"

Group 3: "The (ordinary) laymen.
I: Access to reliable information (critical) + Common reference + User-friendly
B: If answers are not found + Poor (or wrong) answers + User-un-friendly
E: Time saving: Known and consistent reference + High-quality content"

Group 4: "The linguists. **I:** Access to new fields + Clarification on new or overlapping terminology **B:** Diagrams (what is a box and line) + Complex domains are hard to simplify + Risk of oversimplifications **E:** Time saving: Consistency in term use + Ease knowledge dissemination"
Gruppe 1: De brugervenlige I: Hvis ikke noget netværk/gangen (præcision) + sikre den rigtige term B: Hvor/at den eksisterer + bruger den for sjældent/glemmer + erkende behov E: Overblik skattelovgivning (visuelt) + enighed om hvad vi (ikke) taler om + redskab for tekst-folk i KOM (fx Kia) (formuleringer)
Gruppe 2: De skattefaglige I: Kunne sikre fagtermer bruges korrekt (kritisk) + forklaringer til udenforstående (kritisk) + skabe overblik (bakke, zoome) (alvorlig B: tillid (dækkende (inkl. Undtagelser) og ajourført) + tidsrøver (nørder bliver grebet) og dermed for meget information E: Oplæring/Introduktion + Ensartethed (flere (synonyme) ord for det samme: lov, presse, branche) + Fælles reference (fagligheder i samme system) > Tidsbesparende
Gruppe 3: De almindelige I: pålidelig og troværdig information (kritisk) + krav om fælles reference (undgå diskussion om Gyldendal eller Oxford) + brugervenligt B: Hvis der ikke er et svar (ikke fyldestgørende, finder ikke svar) (ikke åbnes, før færdig) + Dårlige svar (svar man ikke kan acceptere) + brugervenlig E: tidsbesparende (hvor + ensartet reference (enighed om validitet) + klarhed) + høj kvalitet på indhold og forståelse
Gruppe 4: De fagsproglige I: Nyt fagområde forkromet overblik + nye smarte ord (bare gamle) B: Kasser og streger (hvad er det?) + komplekse domæner kan være svære at simplificere + risiko for oversimplificering E: Konsekvens i brug af termer + Vidensformidling + Tidsbesparelse ca. 10 min

Bibliography

Authors

Adamska-Sałaciak, A. (2010). Examining equivalence. *International Journal of Lexicography 23* (4), 387-409.

Al-Sayed, R., & Ahmad, K. (2006). Knowledge Sharing in a Community of Practice: a Text-Based Approach in Emergent Domains. In *Proceedings of the 2nd International Conference on Intellectual Capital, Knowledge Management and Organisational Learning (ICICKM 2005)* (p. 33). Academic Conferences Limited.

Alexander, P. A. (1992). Domain knowledge – evolving themes and emerging concerns. *Educational Psychologist, 27 (*1), 33-51.

Andreoni, J., Erard, B. & Feinstein, J. (1998). Tax compliance. *Journal of Economic Literature 36*, 818–860.

Appel, M., & Kronberger, N. (2012). Stereotypes and the achievement gap: Stereotype threat prior to test taking. *Educational Psychology Review, 24*(4), 609-635.

Balling, L. W. (2008). *Morphological Effects in Danish Auditory Word Recognition* (Doctoral dissertation, Aarhus Universitet Aarhus University, [Enhedsstruktur før 1.7. 2011] Aarhus University, Det Humanistiske FakultetFaculty of Humanities, Afdeling for Engelsk Department of English).

Balling, L. W., Hvelplund, K. T. & Sjørup, A. C. (2014) Evidence of Parallel Processing During Translation. In: *Meta 59* (2), 234-259.

Bednarik, R., Kinnunen, T., Mihaila, A., & Fränti, P. (2005). Eye-movements as a biometric. In *Image analysis* (pp. 780-789). Springer Berlin Heidelberg.

Bergenholtz, H. & Gouws, R. H. (2010). A Functional Approach to the Choice between Descriptive, Prescriptive and Proscriptive lexicography. *Lexikos 20*, 26-51.

Bergenholtz, H. & Gouws, R. H. (2012). What is Lexicography? *Lexikos 22*, 31-42.

Bergenholtz, H., & Kaufmann, U. (1997). Terminography and Lexicography. A Critical Survey of Dictionaries from a Single Specialised Field. *Hermes, Journal of Linguistics*.

Bergenholtz, H., & Tarp, S. (2010). LSP lexicography or terminography? The lexicographer's point of view. *FUERTES-OLIVEIRA, PA Specialised dictionaries for learners. Berlin: De Gruyter,* 27-37.

Bilalić, M., McLeod, P., & Gobet, F. (2008). Inflexibility of experts—Reality or myth? Quantifying the Einstellung effect in chess masters. *Cognitive Psychology, 56*(2), 73-102.

Boll, K. (2011). *Taxing assemblages. Laborious and meticulous achievements of tax compliance*. IT University of Copenhagen, The Department, Technologies in Practice. (Doctoral dissertation).

Bonner, S. E., & Walker, P. L. (1994). The effects of instruction and experience on the acquisition of auditing knowledge. *Accounting Review,* 157-178.

Borst, W. N. (1997). *Construction of engineering ontologies for knowledge sharing and reuse*. Universiteit Twente.

Boyle, C., & Encarnacion, A. O. (1998). MetaDoc: an adaptive hypertext reading system. In *Adaptive Hypertext and Hypermedia* (pp. 71-89). Springer Netherlands.

Brewster, C. & O'Hara, K. (2007). Knowledge representation with ontologies: Present challenges—Future possibilities. *International Journal Human-Computer Studies 65,* 563-568.

Britton, B. K., Stimson, M., Stennett, B., & Gülgöz, S. (1998). Learning from instructional text: Test of an individual-differences model. *Journal of Educational Psychology, 90*(3), 476.

Brünken, R., Plass, J. L., & Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual-task methodology: Auditory load and modality effects. *Instructional Science, 32*(1-2), 115-132.

Bundesen, C. & Habekost, T. (2008). Principles of visual attention: Linking mind and brain. Oxford University Press.

Cabré, M. T. (1999). Terminology: Theory, Methods, and Applications. John Benjamins Publishing, 1999.

Cabré, M. T. (2003). Theories of terminology: Their description, prescription and explanation. *Terminology, 9*(2), 163-199.

Caruso, V. (2011). Online specialised dictionaries: a critical survey. In *Electronic lexicography in the 21st century: New Applications for New Users: Proceedings of eLex 2011, Bled, 10-12 November 2011* (pp. 66-75).

Charness, N., & Tuffiash, M. (2008). The role of expertise research and human factors in capturing, explaining, and producing superior performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 50*(3), 427-432.

Chein, M., Mugnier, M. L. & Croitoru, M. (2013). Visual reasoning with graph-based mechanisms: the good, the better and the best. *The Knowledge Engineering Review 28* (03), 249-271.

Conati, C., & Merten, C. (2007). Eye-tracking for user modeling in exploratory learning environments: An empirical evaluation. *Knowledge-Based Systems, 20*(6), 557-574.

Corcho, O., Fernández-López, M., & Gómez-Pérez, A. (2003). Methodologies, tools and languages for building ontologies. Where is their meeting point?. *Data & knowledge engineering, 46*(1), 41-64.

Corcho, O., Fernández-López, M., & Gómez-Pérez, A. (2006). Ontological engineering: principles, methods, tools and languages. In *Ontologies for software engineering and software technology* (pp. 1-48). Springer Berlin Heidelberg.

Cuevas, H. M., Fiore, S. M., & Oser, R. L. (2002). Scaffolding cognitive and metacognitive processes in low verbal ability learners: Use of diagrams in computer-based training environments. *Instructional Science, 30*(6), 433-464.

Da Cunha, I., Cabré, M. T., SanJuan, E., Sierra, G., Torres-Moreno, J. M., & Vivaldi, J. (2011). Automatic specialized vs. Non-specialized sentence differentiation. In *Computational Linguistics and Intelligent Text Processing* (pp. 266-276). Springer Berlin Heidelberg.

De Schryver, G. M. (2003). Lexicographers' Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography*, *16*(2), 143-199.

Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research, 6*(2), 80-88.

Dunbar, K., & Blanchette, I. (2001). The in vivo/in vitro approach to cognition: The case of analogy. *Trends in cognitive sciences*, *5*(8), 334-339.

Fadde, P. J. (2009). Instructional design for advanced learners: Training recognition skills to hasten expertise. *Educational Technology Research and Development*, *57*(3), 359-376.

Fellbaum, C. (2014). Large-Scale Lexicography in the Digital Age. *International Journal of Lexicography*, ecu018.

Fernández-Silva, S., Freixa, J., & Cabré, M. T. (2011). A proposed method for analysing the dynamics of cognition through term variation. *Terminology*, *17*(1), 49-74.

Fischer, R. (1998). Public relations problem solving: Heuristics and expertise. *Journal of Public Relations Research, 10*(2), 137-153.

Flahive, A., Taniar, D., Rahayu, W. & Apduhan, B. O. (2011). Ontology expansion: appending with extracted sub-ontology. *Logic Journal of IGPL 19* (5), 618-647.

Foster, A., & Ford, N. (2003). Serendipity and information seeking: an empirical study. *Journal of Documentation, 59*(3), 321-340.

Frappier, M., Meynell, L. & Brown, J. R. (2013). Special issue on visual representations and reasoning. *The Knowledge Engineering Review 28* (03), 231-236.

Freixa, J. (2006). Causes of denominative variation in terminology: A typology proposal. *Terminology 12* (1), 51–77.

Friege, G., & Lind, G. (2006). Types and qualities of knowledge and their relations to problem solving in physics. *International Journal of Science and Mathematics Education, 4*(3), 437-465.

Fuertes-Olivera, P. A. (2011). Online dictionaries for assisting translators of LSP texts: the Accounting Dictionaries. *International Journal of Lexicography*, ecr034.

Glaser, R. (1989). Expertise and learning: How do we think about instructional processes now that we have discovered knowledge structures. *Complex information processing: The impact of Herbert A. Simon*, 269-282.

Goldberg, J. H. (2014). Measuring Software Screen Complexity: Relating Eye Tracking, Emotional Valence, and Subjective Ratings. *International Journal of Human-Computer Interaction, 30*(7), 518-532.

Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., & Wichansky, A. M. (2002, March). Eye tracking in web search tasks: design implications. In *Proceedings of the 2002 symposium on Eye tracking research & applications* (pp. 51-58). ACM.

Gouws, R. H. (2012). Theoretical lexicography and the International Journal of Lexicography. *International Journal of Lexicography 25* (4), 450-463.

Grabar, N., Hamon, T., & Bodenreider, O. (2012). Ontologies and terminologies: Continuum or dichotomy?. *Applied Ontology, 7*(4), 375-386.

Graves, B. (1996). The study of literary expertise as a research strategy. *Poetics, 23*(6), 385-403.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition, 5*(2), 199-220.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing?. *International journal of human-computer studies, 43*(5), 907-928.

Guarino, N. (1995). Formal ontology, conceptual analysis and knowledge representation. *International journal of human-computer studies, 43*(5), 625-640.

Göpferich, S. (2010). The translation of instructive texts from a cognitive perspective: novices and professionals compared. *New Approaches in Translation Process Research, 39*, 5.

Hart, S. G. (2006, October). NASA-task load index (NASA-TLX); 20 years later. *In Proceedings of the human factors and ergonomics society annual meeting* (Vol. 50, No. 9, pp. 904-908). Sage Publications.

Hart, S.G. & Staveland, L.E. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in psychology, 52*, 139-183.

Hertzum, M. (2010). Images of usability. *Intl. Journal of Human–Computer Interaction, 26*(6), 567-600.

Hofstede, G. (1984). *Culture's consequences: International differences in work-related values* (Vol. 5). sage.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures.* Oxford University Press.

ISO (1998). *Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability (9241-11).* Geneva.

ISO (2000). *Terminology Work - Vocabulary. Part 1: Theory and application (1087-1).* Geneva.

ISO (2009). *Terminology work - Principles and methods (704)*. Geneva.

Jean, S., Pierra, G. & Ait-Ameur, Y. (2007). Domain ontologies: A database-oriented analysis. In *Web Information Systems and Technologies*, pp. 238-254. Springer Berlin Heidelberg.

Jensen, K. T. (2009). Indicators of text complexity. *Behind the mind: Methods, models and results in translation process research, 37*, 61.

Jones, D. M., & Paton, R. C. (1999). Toward principles for the representation of hierarchical knowledge in formal ontologies. *Data & Knowledge Engineering, 31* (2), 99-113.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. Psychological review, 87(4), 329.

Kalyuga, S., & Sweller, J. (2004). Measuring Knowledge to Optimize Cognitive Load Factors During Instruction. *Journal of educational psychology, 96*(3), 558.

Katifori, A., Halatsis, C., Lepouras, G., Vassilakis, C. & Giannopoulou, E. (2007). Ontology visualization methods - a survey. *ACM Computing Surveys (CSUR), 39* (4), 10.

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a randomized tax audit experiment in Denmark. *Econometrica 79* (3), 651-692.

Kobsa, A., Koenemann, J., & Pohl, W. (2001). Personalised hypermedia presentation techniques for improving online customer relationships. *The knowledge engineering review, 16*(02), 111-155.

Koplenig, A. (2011). Understanding How Users Evaluate Innovative Features of Online Dictionaries – An Experimental Approach. In *Proceedings of eLex 2011*.

Kuhn, D. (2000). Metacognitive development. *Current directions in psychological science, 9*(5), 178-181.

Kwak, N. (1999). Revisiting the Knowledge Gap Hypothesis Education, Motivation, and Media Use. *Communication Research, 26*(4), 385-413.

Kwary, D. A. (2012). Adaptive hypermedia and user-oriented data for online dictionaries: A case study on an English dictionary of finance for Indonesian students. *International Journal of Lexicography*, 25(1), 30-49.

L'Homme, M. C. (2003). Capturing the lexical structure in special subject fields with verbs and verbal derivatives. A model for specialized lexicography. *International Journal of Lexicography, 16*(4), 403-422.

L'Homme, M. C. (2014). Why Lexical Semantics is Important for E-Lexicography and Why it is Equally Important to Hide its Formal Representations from Users of Dictionaries. *International Journal of Lexicography*, 27(4), 360-377.

Lachner, A., Gurlitt, J., & Nückles, M. (2012). A graph-oriented approach to measuring expertise—detecting structural differences between experts and intermediates. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 653-658).

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive science, 11*(1), 65-100.

Lee, M. D., Steyvers, M., De Young, M., & Miller, B. J. (2011). A model-based approach to measuring expertise in ranking tasks. In *Proceedings of the 33rd annual conference of the cognitive science society*.

Lew, R. (2010). Multimodal lexicography: The representation of meaning in electronic dictionaries. *Lexikos, 20*(1), 290-306.

Lew, R., & De Schryver, G. M. (2014). Dictionary users in the digital revolution. *International Journal of Lexicography*, ecu011.

Loboda, T. D., & Brusilovsky, P. (2010). User-adaptive explanatory program visualization: evaluation and insights from eye movements. *User Modeling and User-Adapted Interaction, 20*(3), 191-226.

Madsen, B. N. (2006). Terminological ontologies: Applications and principles. Semantic Systems: *From Visions to Applications. Proceedings of the Semantics 2006. Österreichische Computer Gesellschaft*, 271-282.

Madsen, B. N., Thomsen, H. E., Halskov, J. & Lassen, T. (2010). Automatic Ontology Construction for a National Term Bank. In *Terminology and Knowledge Engineering Conference 2010*, pp. 502-533.

Madsen, B. N., Thomsen, H. E., & Vikner, C. (2004). Comparison of Principles Applying to Domain-Specific versus General Ontologies. In *OntoLex* (pp. 90-95).

Mahalingam, K., & Huhns, M. N. (1998). Representing and using Ontologies. *Center for Information Technology, Thesis, University of South Carolina*.

Mason, A., & Singh, C. (2011). Assessing expertise in introductory physics using categorization task. *Physical Review Special Topics-Physics Education Research, 7*(2), 020110.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist, 38*(1), 43-52.

Mertens, D. M. (2012). What comes first? The paradigm or the approach? *Journal of Mixed Methods Research, 6*(4), 255-257.

Mourier, L., & Vesterli, B. (2009). KNowaboUT: A Work-In-Progress Study on How to Fruitfully Combine the Theories of Terminology and Lexicography. *Nordterm 2009*.

Müller-Spitzer, C., Wolfer, S., & Koplenig, A. (2015). Observing online dictionary users: Studies using wiktionary log files. *International Journal of Lexicography*, *28*(1), 1-26.

Nielsen, L. (2004). Engaging personas and narrative scenarios. *PhD series, 17.*

Nkwenti-Azeh, B. (1994, March). New trends in terminology processing and implications for practical translation. In *Aslib proceedings* (Vol. 46, No. 3, pp. 67-74). MCB UP Ltd.

Noy, N. F. (2004). Semantic integration: a survey of ontology-based approaches. *ACM Sigmod Record*, *33*(4), 65-70.

Ormerod, T. C., & Ball, L. J. (2008). Cognitive psychology. *Handbook of qualitative research in psychology*, 554-574.

Osman, M. E., & Hannafin, M. J. (1992). Metacognition research and theory: Analysis and implications for instructional design. *Educational Technology Research and Development, 40*(2), 83-99.

Paavola, S., Lipponen, L., & Hakkarainen, K. (2004). Models of innovative knowledge communities and three metaphors of learning. *Review of educational research, 74*(4), 557-576.

Pacharapha, T., & Vathanophas Ractham, V. (2012). Knowledge acquisition: the roles of perceived value of knowledge content and source. *Journal of Knowledge Management, 16*(5), 724-739.

Polguére, A. (2014). From Writing Dictionaries to Weaving Lexical Networks. *International Journal of Lexicography*, ecu017.

Rayner, K. (2009). Eye movements in reading: Models and data. *Journal of eye movement research, 2*(5), 1.

Rayner, K., Juhasz, B. J., & Pollatsek, A. (2005). Eye movements during reading. *The science of reading: A handbook*, 79-97.

Rikers, R. M., & Paas, F. (2005). Recent advances in expertise research. *Applied Cognitive Psychology, 19*(2), 145-149.

Roche, C. (2006). Lexical and conceptual structures in ontology. *Advances in applied artificial intelligence, proceedings.* Book Series: Lecture notes in computer science, vol. 4031, pp. 1034-1041.

Roche, C., Calberg-Challot, M., Damas, L., & Rouard, P. (2009). Ontoterminology: A new paradigm for terminology. In *International Conference on Knowledge Engineering and Ontology Development* (pp. 321-326).

Rogers, M. (2004). Multidimensionality in concepts systems: A bilingual textual perspective. *Terminology 10*(2), 215–240.

Ruben, J. & Chisnell, D. (2008) Handbook of Usability Testing : How to Plan, Design, and Conduct Effective Tests. Wiley. Hoboken, NJ, USA.

Rylova, A. (2011). Adapting the dictionary entry structure and DWS configuration for creating a dictionary aimed to be published on paper, online and as electronic dictionary software for PC and mobile. *In Electronic lexicography in the 21st century: New Applications for New Users: Proceedings of eLex 2011, Bled, 10-12 November 2011* (pp. 247-250).

Rusanen, A. M., & Pöyhönen, S. (2013). Concepts in change. *Science & Education*, *22*(6), 1389-1403.

Sager, J. C. (1990): A practical course in terminology processing. John Benjamin Publishing.

Sandrini, P. (1999). Legal Terminology. Some Aspects for a New Methodology. *Hermes Journal of Linguistics*, *22*, 101-112.

Schnotz, W., & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology Research and Development*, *53*(3), 47-58.

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and instruction*, *13*(2), 227-237.

Simonsen, J. & Robertson, T. (2013). Routledge International Handbook of Participatory Design.

Solomon, Medin & Lynch, 1999. Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts?. *Annual review of psychology*, *51*(1), 121-147.

Studer, R., Benjamins, V.R., Fensel, D. (1998). Knowledge Engineering: Principles and methods. *Data and Knowledge Engineering 25* (1-2), 161-197.

Tarp, S. (1998). Leksikografien på egne ben. Fordelingsstrukturer og byggedele i et brugerorienteret perspektiv. Hermes 21: 121-137.

Tarp, S. (2000). Tarp, S. (2000). Theoretical challenges to practical specialised lexicography. *Lexikos*, *10*(1).

Tarp, S. (2010). Functions of specialised learners' dictionaries. *Specialised Dictionaries for Learners*, Berlin y Nueva York: De Gruyter, 39-53.

Thellefsen, T. L. & Thellefsen, M. (2004). Semiotik og vidensorganisering. In Thellefsen, T. L. & Dinesen, A. M. *Semiotiske undersøgelser*. Reitzel.

Thomsen, H. E. (2012a). Term Banks. *The Encyclopedia of Applied Linguistics.*

Thomsen, H. E. (2012b). The DanTerm Project : Automatic Knowledge Extraction and Knowledge Structuring for a National Term Bank. Poster presented at *The 6th Terminology Summit*, Oslo, Norway.

Tono, Y. (2011). Application of eye-tracking in EFL learners’ dictionary look-up process research. *International Journal of Lexicography*, *24*(1), 124-153.

Verlinde, S., Leroyer, P., & Binon, J. (2009). Search and you will find. From stand-alone lexicographic tools to user driven task and problem-oriented multifunctional leximats. *International Journal of Lexicography*, ecp029.

Wüster, E. 1968. *The Machine Tool. An interlingual dictionary of basic concepts*. London: Technical Press.

Wüster, E. 1974. Die allgemeine Terminologielehre — Ein Grenzgebiet zwischen Sprachwissenschaft, Logik, Ontologie, Informatik und den Sachwissenschaften. *Linguistics 119*: 61–106.

Online ressources

Baayen, R. H. (2011). *languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics." R package version 1.4* Retrieved September 2, 2013, from http://CRAN.R-project.org/ package=languageR.

Bates, D. Maechler, M., Bolker, B. & Walker, S. (2013). *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-5*. Retrieved September 2, 2013, from http://CRAN.R-project.org/package=lme4.

DanNet (2015). *Lexical ontology for Danish*. Retrieved February 15, 2015, from http://wordnet.dk/.

DanTermBank (2015a). *Knowledge dissemination.* Retrieved April 11, 2015, from http://dantermbank.cbs.dk/dtb_uk/about_the_project/knowledge_ dissemination.

DanTermBank (2015b). *The ontology of ontologies for OntoRobot*. Forthcoming, on http://dantermbank.cbs.dk/dtb_uk/publications.

DanTermBank (2015c). *The project overview diagram.* Retrieved April 11, 2015, from from http://dantermbank.cbs.dk/dtb_uk/about_the_project

DanTermBank (2015d). *User Scenarios*. Retrieved April 11, 2015, from http://dantermbank.cbs.dk/dtb_uk/about_the_project/knowledge_disseminati on/user_scenarios.

EuroStat (2012). *Statistic glossary from the statistical office of the European Union*. Retrieved November 20, 2012, from http://ec.europa.eu/eurostat/.

EuroStat (2014). *Publications. Taxation Trends in the EU 2012.* Retrieved September 18, 2014 from http://ec.europa.eu/taxation_customs/taxation /gen_info/economic_analysis/tax_structures/index_en.htm.

Gyldendal (2015a). *National encyclopaedic reference work.* Retrieved February 15, 2015, from http://denstoredanske.dk.

Gyldendal (2015b). *English-Danish dictionary*. Retrieved February 15, 2015, from http://www.ordbog.gyldendal.com.

IATE (2015). The EU's multilingual term base. Retrieved February 15, 2015, from http://iate.europa.eu/SearchByQueryLoad.do?method=load.

Ministry of Finance (2014). *Publications*. *The finance and appropriation act regarding 2012*. Retrieved September 18, 2014, from http://www.fm.dk/publikationer.

OECD (2014). *Centre for Tax Policy and Administration. Revenue Statistics 1965-2010*. Retrieved September 18, 2014, from http://www.oecd.org/ctp/tax-policy/revenue-statistics-19963726.htm.

OECD (2015). *Glossary of tax terms*. Retrieved February 15, 2015, from http://www.oecd.org/ctp/glossaryoftaxterms.htm.

Ordbogen (2015). *English-Danish-English dictionary*. Retrieved February 15, 2015, from http://www.ordbogen.com.

Oxford Dictionaries (2015). *Online dictionary of British and world English by Oxford University Press*. Retrieved January 30, 2015, from http:// http://www.oxforddictionaries.com/.

R Core Team (2013). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria*. Retrieved September 2, 2013, from http://www.R-project.org/.

Retsinformation (2015a). *Danish legal data base for the Ministry of Legal affairs*. Retrieved February 15, 2015, from https://www.retsinformation.dk/Forms/R0310.aspx?res=1.

Retsinformation (2015b). *Danish legal data base for the Ministry of Taxation*. Retrieved February 15, 2015, from https://www.retsinformation.dk/Forms/R0310.aspx?res=14.

Rikstermbanken (2015). *Sweden's national term bank*. Retrieved February 15, 2015, from http://www.rikstermbanken.se/.

SKAT (2014). *Legal instruction data base*. Retrieved September 18, 2014, from http://skat.dk/SKAT.aspx?oId=124.

Språkrådet (2015). *Termwiki for terminology and specialized language in Norway*. Retrieved February 15, 2015, from http://termwiki.sprakradet.no/.

SSI (2015). *Health care terminology*. Retrieved February 15, 2015, from http://begrebsbasen.sst.dk/.

Statistics Denmark (2014). *Taxes and Duties 2012*. Retrieved September 18, 2014, from http://www.dst.dk/da/Statistik/emner/skatter-og-afgifter/skatter-og-afgifter.aspx.

Tax Policy Center (2015). *Tax glossary*. Retrieved February 15, 2015, from http://www.taxpolicycenter.org/.

University of Vaasa (2015). *University glossaries*. Retrieved February 15, 2015, from http://www.uva.fi/.

Usability First (2015). *Glossary consisting of popular terms pertaining to User-Centered Design by Foraker Labs*. Retrieved February 15, 2015, from http://www.usabilityfirst.com/glossary/.

WIPO (2015). *Multilingual terminology portal gives access to scientific and technical terms derived from patent documents*. Retrieved February 15, 2015, from http://www.wipo.int/.

World Bank (2015). *Glossary related to projects and operations*. Retrieved February 15, 2015, from http://www.worldbank.org/.

TITLER I PH.D.SERIEN:

2004

1. Martin Grieger
Internet-based Electronic Marketplaces and Supply Chain Management

2. Thomas Basbøll
LIKENESS
A Philosophical Investigation

3. Morten Knudsen
Beslutningens vaklen
En systemteoretisk analyse of moderniseringen af et amtskommunalt sundhedsvæsen 1980-2000

4. Lars Bo Jeppesen
Organizing Consumer Innovation
A product development strategy that is based on online communities and allows some firms to benefit from a distributed process of innovation by consumers

5. Barbara Dragsted
SEGMENTATION IN TRANSLATION AND TRANSLATION MEMORY SYSTEMS
An empirical investigation of cognitive segmentation and effects of integrating a TM system into the translation process

6. Jeanet Hardis
Sociale partnerskaber
Et socialkonstruktivistisk casestudie af partnerskabsaktørers virkelighedsopfattelse mellem identitet og legitimitet

7. Henriette Hallberg Thygesen
System Dynamics in Action

8. Carsten Mejer Plath
Strategisk Økonomistyring

9. Annemette Kjærgaard
Knowledge Management as Internal Corporate Venturing
– a Field Study of the Rise and Fall of a Bottom-Up Process

10. Knut Arne Hovdal
De profesjonelle i endring
Norsk ph.d., ej til salg gennem Samfundslitteratur

11. Søren Jeppesen
Environmental Practices and Greening Strategies in Small Manufacturing Enterprises in South Africa
– A Critical Realist Approach

12. Lars Frode Frederiksen
Industriel forskningsledelse
– på sporet af mønstre og samarbejde i danske forskningsintensive virksomheder

13. Martin Jes Iversen
The Governance of GN Great Nordic
– in an age of strategic and structural transitions 1939-1988

14. Lars Pynt Andersen
The Rhetorical Strategies of Danish TV Advertising
A study of the first fifteen years with special emphasis on genre and irony

15. Jakob Rasmussen
Business Perspectives on E-learning

16. Sof Thrane
The Social and Economic Dynamics of Networks
– a Weberian Analysis of Three Formalised Horizontal Networks

17. Lene Nielsen
Engaging Personas and Narrative Scenarios – a study on how a user-centered approach influenced the perception of the design process in the e-business group at AstraZeneca

18. S.J Valstad
Organisationsidentitet
Norsk ph.d., ej til salg gennem Samfundslitteratur

19. Thomas Lyse Hansen
Six Essays on Pricing and Weather risk in Energy Markets

20. Sabine Madsen
Emerging Methods – An Interpretive Study of ISD Methods in Practice

21. Evis Sinani
The Impact of Foreign Direct Investment on Efficiency, Productivity Growth and Trade: An Empirical Investigation

22. Bent Meier Sørensen
Making Events Work Or, How to Multiply Your Crisis

23. Pernille Schnoor
Brand Ethos
Om troværdige brand- og virksomhedsidentiteter i et retorisk og diskursteoretisk perspektiv

24. Sidsel Fabech
Von welchem Österreich ist hier die Rede?
Diskursive forhandlinger og magtkampe mellem rivaliserende nationale identitetskonstruktioner i østrigske pressediskurser

25. Klavs Odgaard Christensen
Sprogpolitik og identitetsdannelse i flersprogede forbundsstater
Et komparativt studie af Schweiz og Canada

26. Dana B. Minbaeva
Human Resource Practices and Knowledge Transfer in Multinational Corporations

27. Holger Højlund
Markedets politiske fornuft
Et studie af velfærdens organisering i perioden 1990-2003

28. Christine Mølgaard Frandsen
A.s erfaring
Om mellemværendets praktik i en transformation af mennesket og subjektiviteten

29. Sine Nørholm Just
The Constitution of Meaning
– A Meaningful Constitution?
Legitimacy, identity, and public opinion in the debate on the future of Europe

2005

1. Claus J. Varnes
Managing product innovation through rules – The role of formal and structured methods in product development

2. Helle Hedegaard Hein
Mellem konflikt og konsensus
– Dialogudvikling på hospitalsklinikker

3. Axel Rosenø
Customer Value Driven Product Innovation – A Study of Market Learning in New Product Development

4. Søren Buhl Pedersen
Making space
An outline of place branding

5. Camilla Funck Ellehave
Differences that Matter
An analysis of practices of gender and organizing in contemporary workplaces

6. Rigmor Madeleine Lond
Styring af kommunale forvaltninger

7. Mette Aagaard Andreassen
Supply Chain versus Supply Chain
Benchmarking as a Means to Managing Supply Chains

8. Caroline Aggestam-Pontoppidan
From an idea to a standard
The UN and the global governance of accountants' competence

9. Norsk ph.d.

10. Vivienne Heng Ker-ni
An Experimental Field Study on the

Effectiveness of Grocer Media Advertising
Measuring Ad Recall and Recognition, Purchase Intentions and Short-Term Sales

11. Allan Mortensen
Essays on the Pricing of Corporate Bonds and Credit Derivatives

12. Remo Stefano Chiari
Figure che fanno conoscere
Itinerario sull'idea del valore cognitivo e espressivo della metafora e di altri tropi da Aristotele e da Vico fino al cognitivismo contemporaneo

13. Anders McIlquham-Schmidt
Strategic Planning and Corporate Performance
An integrative research review and a meta-analysis of the strategic planning and corporate performance literature from 1956 to 2003

14. Jens Geersbro
The TDF – PMI Case
Making Sense of the Dynamics of Business Relationships and Networks

15 Mette Andersen
Corporate Social Responsibility in Global Supply Chains
Understanding the uniqueness of firm behaviour

16. Eva Boxenbaum
Institutional Genesis: Micro – Dynamic Foundations of Institutional Change

17. Peter Lund-Thomsen
Capacity Development, Environmental Justice NGOs, and Governance: The Case of South Africa

18. Signe Jarlov
Konstruktioner af offentlig ledelse

19. Lars Stæhr Jensen
Vocabulary Knowledge and Listening Comprehension in English as a Foreign Language
An empirical study employing data elicited from Danish EFL learners

20. Christian Nielsen
Essays on Business Reporting
Production and consumption of strategic information in the market for information

21. Marianne Thejls Fischer
Egos and Ethics of Management Consultants

22. Annie Bekke Kjær
Performance management i Proces-innovation
– belyst i et social-konstruktivistisk perspektiv

23. Suzanne Dee Pedersen
GENTAGELSENS METAMORFOSE
Om organisering af den kreative gøren i den kunstneriske arbejdspraksis

24. Benedikte Dorte Rosenbrink
Revenue Management
Økonomiske, konkurrencemæssige & organisatoriske konsekvenser

25. Thomas Riise Johansen
Written Accounts and Verbal Accounts
The Danish Case of Accounting and Accountability to Employees

26. Ann Fogelgren-Pedersen
The Mobile Internet: Pioneering Users' Adoption Decisions

27. Birgitte Rasmussen
Ledelse i fællesskab – de tillidsvalgtes fornyende rolle

28. Gitte Thit Nielsen
Remerger
– skabende ledelseskræfter i fusion og opkøb

29. Carmine Gioia
A MICROECONOMETRIC ANALYSIS OF MERGERS AND ACQUISITIONS

30. Ole Hinz
Den effektive forandringsleder: pilot, pædagog eller politiker?
Et studie i arbejdslederes meningstilskrivninger i forbindelse med vellykket gennemførelse af ledelsesinitierede forandringsprojekter

31. Kjell-Åge Gotvassli
Et praksisbasert perspektiv på dynamiske
læringsnettverk i toppidretten
Norsk ph.d., ej til salg gennem Samfundslitteratur

32. Henriette Langstrup Nielsen
Linking Healthcare
An inquiry into the changing performances of web-based technology for asthma monitoring

33. Karin Tweddell Levinsen
Virtuel Uddannelsespraksis
Master i IKT og Læring – et casestudie i hvordan proaktiv proceshåndtering kan forbedre praksis i virtuelle læringsmiljøer

34. Anika Liversage
Finding a Path
Labour Market Life Stories of Immigrant Professionals

35. Kasper Elmquist Jørgensen
Studier i samspillet mellem stat og erhvervsliv i Danmark under 1. verdenskrig

36. Finn Janning
A DIFFERENT STORY
Seduction, Conquest and Discovery

37. Patricia Ann Plackett
Strategic Management of the Radical Innovation Process
Leveraging Social Capital for Market Uncertainty Management

2006

1. Christian Vintergaard
Early Phases of Corporate Venturing

2. Niels Rom-Poulsen
Essays in Computational Finance

3. Tina Brandt Husman
Organisational Capabilities, Competitive Advantage & Project-Based Organisations
The Case of Advertising and Creative Good Production

4. Mette Rosenkrands Johansen
Practice at the top
– how top managers mobilise and use non-financial performance measures

5. Eva Parum
Corporate governance som strategisk kommunikations- og ledelsesværktøj

6. Susan Aagaard Petersen
Culture's Influence on Performance Management: The Case of a Danish Company in China

7. Thomas Nicolai Pedersen
The Discursive Constitution of Organizational Governance – Between unity and differentiation
The Case of the governance of environmental risks by World Bank environmental staff

8. Cynthia Selin
Volatile Visions: Transactons in Anticipatory Knowledge

9. Jesper Banghøj
Financial Accounting Information and Compensation in Danish Companies

10. Mikkel Lucas Overby
Strategic Alliances in Emerging High-Tech Markets: What's the Difference and does it Matter?

11. Tine Aage
External Information Acquisition of Industrial Districts and the Impact of Different Knowledge Creation Dimensions

A case study of the Fashion and Design Branch of the Industrial District of Montebelluna, NE Italy

12. Mikkel Flyverbom
Making the Global Information Society Governable
On the Governmentality of Multi-Stakeholder Networks

13. Anette Grønning
Personen bag
Tilstedevær i e-mail som inter-aktionsform mellem kunde og med-arbejder i dansk forsikringskontekst

14. Jørn Helder
One Company – One Language?
The NN-case

15. Lars Bjerregaard Mikkelsen
Differing perceptions of customer value
Development and application of a tool for mapping perceptions of customer value at both ends of customer-suppli-er dyads in industrial markets

16. Lise Granerud
Exploring Learning
Technological learning within small manufacturers in South Africa

17. Esben Rahbek Pedersen
Between Hopes and Realities: Reflections on the Promises and Practices of Corporate Social Responsibility (CSR)

18. Ramona Samson
The Cultural Integration Model and European Transformation.
The Case of Romania

2007

1. Jakob Vestergaard
Discipline in The Global Economy
Panopticism and the Post-Washington Consensus

2. Heidi Lund Hansen
Spaces for learning and working
A qualitative study of change of work, management, vehicles of power and social practices in open offices

3. Sudhanshu Rai
Exploring the internal dynamics of software development teams during user analysis
A tension enabled Institutionalization Model; "Where process becomes the objective"

4. Norsk ph.d.
Ej til salg gennem Samfundslitteratur

5. Serden Ozcan
EXPLORING HETEROGENEITY IN ORGANIZATIONAL ACTIONS AND OUTCOMES
A Behavioural Perspective

6. Kim Sundtoft Hald
Inter-organizational Performance Measurement and Management in Action
– An Ethnography on the Construction of Management, Identity and Relationships

7. Tobias Lindeberg
Evaluative Technologies
Quality and the Multiplicity of Performance

8. Merete Wedell-Wedellsborg
Den globale soldat
Identitetsdannelse og identitetsledelse i multinationale militære organisatio-ner

9. Lars Frederiksen
Open Innovation Business Models
Innovation in firm-hosted online user communities and inter-firm project ventures in the music industry
– A collection of essays

10. Jonas Gabrielsen
Retorisk toposlære – fra statisk 'sted' til persuasiv aktivitet

11. Christian Moldt-Jørgensen
Fra meningsløs til meningsfuld evaluering.
Anvendelsen af studentertilfredsheds-målinger på de korte og mellemlange videregående uddannelser set fra et psykodynamisk systemperspektiv

12. Ping Gao
Extending the application of actor-network theory
Cases of innovation in the tele-communications industry

13. Peter Mejlby
Frihed og fængsel, en del af den samme drøm?
Et phronetisk baseret casestudie af frigørelsens og kontrollens sam-eksistens i værdibaseret ledelse!

14. Kristina Birch
Statistical Modelling in Marketing

15. Signe Poulsen
Sense and sensibility:
The language of emotional appeals in insurance marketing

16. Anders Bjerre Trolle
Essays on derivatives pricing and dyna-mic asset allocation

17. Peter Feldhütter
Empirical Studies of Bond and Credit Markets

18. Jens Henrik Eggert Christensen
Default and Recovery Risk Modeling and Estimation

19. Maria Theresa Larsen
Academic Enterprise: A New Mission for Universities or a Contradiction in Terms?
Four papers on the long-term impli-cations of increasing industry involve-ment and commercialization in acade-mia

20. Morten Wellendorf
Postimplementering af teknologi i den offentlige forvaltning
Analyser af en organisations konti-nuerlige arbejde med informations-teknologi

21. Ekaterina Mhaanna
Concept Relations for Terminological Process Analysis

22. Stefan Ring Thorbjørnsen
Forsvaret i forandring
Et studie i officerers kapabiliteter un-der påvirkning af omverdenens foran-dringspres mod øget styring og læring

23. Christa Breum Amhøj
Det selvskabte medlemskab om ma-nagementstaten, dens styringstekno-logier og indbyggere

24. Karoline Bromose
Between Technological Turbulence and Operational Stability
– An empirical case study of corporate venturing in TDC

25. Susanne Justesen
Navigating the Paradoxes of Diversity in Innovation Practice
– A Longitudinal study of six very different innovation processes – in practice

26. Luise Noring Henler
Conceptualising successful supply chain partnerships
– Viewing supply chain partnerships from an organisational culture per-spective

27. Mark Mau
Kampen om telefonen
Det danske telefonvæsen under den tyske besættelse 1940-45

28. Jakob Halskov
The semiautomatic expansion of existing terminological ontologies using knowledge patterns discovered

on the WWW – an implementation and evaluation

29. Gergana Koleva
European Policy Instruments Beyond Networks and Structure: The Innovative Medicines Initiative

30. Christian Geisler Asmussen
Global Strategy and International Diversity: A Double-Edged Sword?

31. Christina Holm-Petersen
Stolthed og fordom
Kultur- og identitetsarbejde ved skabelsen af en ny sengeafdeling gennem fusion

32. Hans Peter Olsen
Hybrid Governance of Standardized States
Causes and Contours of the Global Regulation of Government Auditing

33. Lars Bøge Sørensen
Risk Management in the Supply Chain

34. Peter Aagaard
Det unikkes dynamikker
De institutionelle mulighedsbetingelser bag den individuelle udforskning i professionelt og frivilligt arbejde

35. Yun Mi Antorini
Brand Community Innovation
An Intrinsic Case Study of the Adult Fans of LEGO Community

36. Joachim Lynggaard Boll
Labor Related Corporate Social Performance in Denmark
Organizational and Institutional Perspectives

2008

1. Frederik Christian Vinten
Essays on Private Equity

2. Jesper Clement
Visual Influence of Packaging Design on In-Store Buying Decisions

3. Marius Brostrøm Kousgaard
Tid til kvalitetsmåling?
– Studier af indrulleringsprocesser i forbindelse med introduktionen af kliniske kvalitetsdatabaser i speciallægepraksissektoren

4. Irene Skovgaard Smith
Management Consulting in Action
Value creation and ambiguity in client-consultant relations

5. Anders Rom
Management accounting and integrated information systems
How to exploit the potential for management accounting of information technology

6. Marina Candi
Aesthetic Design as an Element of Service Innovation in New Technology-based Firms

7. Morten Schnack
Teknologi og tværfaglighed
– en analyse af diskussionen omkring indførelse af EPJ på en hospitalsafdeling

8. Helene Balslev Clausen
Juntos pero no revueltos – un estudio sobre emigrantes norteamericanos en un pueblo mexicano

9. Lise Justesen
Kunsten at skrive revisionsrapporter.
En beretning om forvaltningsrevisionens beretninger

10. Michael E. Hansen
The politics of corporate responsibility: CSR and the governance of child labor and core labor rights in the 1990s

11. Anne Roepstorff
Holdning for handling – en etnologisk undersøgelse af Virksomheders Sociale Ansvar/CSR

12. Claus Bajlum
Essays on Credit Risk and Credit Derivatives

13. Anders Bojesen
The Performative Power of Competence – an Inquiry into Subjectivity and Social Technologies at Work

14. Satu Reijonen
Green and Fragile
A Study on Markets and the Natural Environment

15. Ilduara Busta
Corporate Governance in Banking
A European Study

16. Kristian Anders Hvass
A Boolean Analysis Predicting Industry Change: Innovation, Imitation & Business Models
The Winning Hybrid: A case study of isomorphism in the airline industry

17. Trine Paludan
De uvidende og de udviklingsparate
Identitet som mulighed og restriktion blandt fabriksarbejdere på det aftayloriserede fabriksgulv

18. Kristian Jakobsen
Foreign market entry in transition economies: Entry timing and mode choice

19. Jakob Elming
Syntactic reordering in statistical machine translation

20. Lars Brømsøe Termansen
Regional Computable General Equilibrium Models for Denmark
Three papers laying the foundation for regional CGE models with agglomeration characteristics

21. Mia Reinholt
The Motivational Foundations of Knowledge Sharing

22. Frederikke Krogh-Meibom
The Co-Evolution of Institutions and Technology
– A Neo-Institutional Understanding of Change Processes within the Business Press – the Case Study of Financial Times

23. Peter D. Ørberg Jensen
OFFSHORING OF ADVANCED AND HIGH-VALUE TECHNICAL SERVICES: ANTECEDENTS, PROCESS DYNAMICS AND FIRMLEVEL IMPACTS

24. Pham Thi Song Hanh
Functional Upgrading, Relational Capability and Export Performance of Vietnamese Wood Furniture Producers

25. Mads Vangkilde
Why wait?
An Exploration of first-mover advantages among Danish e-grocers through a resource perspective

26. Hubert Buch-Hansen
Rethinking the History of European Level Merger Control
A Critical Political Economy Perspective

2009

1. Vivian Lindhardsen
From Independent Ratings to Communal Ratings: A Study of CWA Raters' Decision-Making Behaviours

2. Guðrið Weihe
Public-Private Partnerships: Meaning and Practice

3. Chris Nøkkentved
Enabling Supply Networks with Collaborative Information Infrastructures
An Empirical Investigation of Business Model Innovation in Supplier Relationship Management

4. Sara Louise Muhr
Wound, Interrupted – On the Vulnerability of Diversity Management

5. Christine Sestoft
Forbrugeradfærd i et Stats- og Livsformsteoretisk perspektiv

6. Michael Pedersen
Tune in, Breakdown, and Reboot: On the production of the stress-fit self-managing employee

7. Salla Lutz
Position and Reposition in Networks – Exemplified by the Transformation of the Danish Pine Furniture Manufacturers

8. Jens Forssbæck
Essays on market discipline in commercial and central banking

9. Tine Murphy
Sense from Silence – A Basis for Organised Action
How do Sensemaking Processes with Minimal Sharing Relate to the Reproduction of Organised Action?

10. Sara Malou Strandvad
Inspirations for a new sociology of art: A sociomaterial study of development processes in the Danish film industry

11. Nicolaas Mouton
On the evolution of social scientific metaphors:
A cognitive-historical enquiry into the divergent trajectories of the idea that collective entities – states and societies, cities and corporations – are biological organisms.

12. Lars Andreas Knutsen
Mobile Data Services:
Shaping of user engagements

13. Nikolaos Theodoros Korfiatis
Information Exchange and Behavior
A Multi-method Inquiry on Online Communities

14. Jens Albæk
Forestillinger om kvalitet og tværfaglighed på sygehuse
– skabelse af forestillinger i læge- og plejegrupperne angående relevans af nye idéer om kvalitetsudvikling gennem tolkningsprocesser

15. Maja Lotz
The Business of Co-Creation – and the Co-Creation of Business

16. Gitte P. Jakobsen
Narrative Construction of Leader Identity in a Leader Development Program Context

17. Dorte Hermansen
"Living the brand" som en brandorienteret dialogisk praxis:
Om udvikling af medarbejdernes brandorienterede dømmekraft

18. Aseem Kinra
Supply Chain (logistics) Environmental Complexity

19. Michael Nørager
How to manage SMEs through the transformation from non innovative to innovative?

20. Kristin Wallevik
Corporate Governance in Family Firms
The Norwegian Maritime Sector

21. Bo Hansen Hansen
Beyond the Process
Enriching Software Process Improvement with Knowledge Management

22. Annemette Skot-Hansen
Franske adjektivisk afledte adverbier, der tager præpositionssyntagmer indledt med præpositionen à som argumenter
En valensgrammatisk undersøgelse

23. Line Gry Knudsen
Collaborative R&D Capabilities
In Search of Micro-Foundations

24. Christian Scheuer
Employers meet employees
Essays on sorting and globalization

25. Rasmus Johnsen
The Great Health of Melancholy
A Study of the Pathologies of Performativity

26. Ha Thi Van Pham
Internationalization, Competitiveness Enhancement and Export Performance of Emerging Market Firms:
Evidence from Vietnam

27. Henriette Balieu
Kontrolbegrebets betydning for kausativalternationen i spansk
En kognitiv-typologisk analyse

2010

1. Yen Tran
Organizing Innovationin Turbulent Fashion Market
Four papers on how fashion firms create and appropriate innovation value

2. Anders Raastrup Kristensen
Metaphysical Labour
Flexibility, Performance and Commitment in Work-Life Management

3. Margrét Sigrún Sigurdardottir
Dependently independent
Co-existence of institutional logics in the recorded music industry

4. Ásta Dis Óladóttir
Internationalization from a small domestic base:
An empirical analysis of Economics and Management

5. Christine Secher
E-deltagelse i praksis – politikernes og forvaltningens medkonstruktion og konsekvenserne heraf

6. Marianne Stang Våland
What we talk about when we talk about space:
End User Participation between Processes of Organizational and Architectural Design

7. Rex Degnegaard
Strategic Change Management
Change Management Challenges in the Danish Police Reform

8. Ulrik Schultz Brix
Værdi i rekruttering – den sikre beslutning
En pragmatisk analyse af perception og synliggørelse af værdi i rekrutterings- og udvælgelsesarbejdet

9. Jan Ole Similä
Kontraktsledelse
Relasjonen mellom virksomhetsledelse og kontraktshåndtering, belyst via fire norske virksomheter

10. Susanne Boch Waldorff
Emerging Organizations: In between local translation, institutional logics and discourse

11. Brian Kane
Performance Talk
Next Generation Management of Organizational Performance

12. Lars Ohnemus
Brand Thrust: Strategic Branding and Shareholder Value
An Empirical Reconciliation of two Critical Concepts

13. Jesper Schlamovitz
Håndtering af usikkerhed i film- og byggeprojekter

14. Tommy Moesby-Jensen
Det faktiske livs forbindtlighed
Førsokratisk informeret, ny-aristotelisk ἦθος-tænkning hos Martin Heidegger

15. Christian Fich
Two Nations Divided by Common Values
French National Habitus and the Rejection of American Power

16. Peter Beyer
Processer, sammenhængskraft og fleksibilitet Et empirisk casestudie af omstillingsforløb i fire virksomheder

17. Adam Buchhorn
Markets of Good Intentions Constructing and Organizing Biogas Markets Amid Fragility and Controversy

18. Cecilie K. Moesby-Jensen
Social læring og fælles praksis Et mixed method studie, der belyser læringskonsekvenser af et lederkursus for et praksisfællesskab af offentlige mellemledere

19. Heidi Boye
Fødevarer og sundhed i senmodernismen – En indsigt i hyggefænomenet og de relaterede fødevarepraksisser

20. Kristine Munkgård Pedersen
Flygtige forbindelser og midlertidige mobiliseringer Om kulturel produktion på Roskilde Festival

21. Oliver Jacob Weber
Causes of Intercompany Harmony in Business Markets – An Empirical Investigation from a Dyad Perspective

22. Susanne Ekman
Authority and Autonomy Paradoxes of Modern Knowledge Work

23. Anette Frey Larsen
Kvalitetsledelse på danske hospitaler – Ledelsernes indflydelse på introduktion og vedligeholdelse af kvalitetsstrategier i det danske sundhedsvæsen

24. Toyoko Sato
Performativity and Discourse: Japanese Advertisements on the Aesthetic Education of Desire

25. Kenneth Brinch Jensen
Identifying the Last Planner System Lean management in the construction industry

26. Javier Busquets
Orchestrating Network Behavior for Innovation

27. Luke Patey
The Power of Resistance: India's National Oil Company and International Activism in Sudan

28. Mette Vedel
Value Creation in Triadic Business Relationships. Interaction, Interconnection and Position

29. Kristian Tørning
Knowledge Management Systems in Practice – A Work Place Study

30. Qingxin Shi
An Empirical Study of Thinking Aloud Usability Testing from a Cultural Perspective

31. Tanja Juul Christiansen
Corporate blogging: Medarbejderes kommunikative handlekraft

32. Malgorzata Ciesielska
Hybrid Organisations. A study of the Open Source – business setting

33. Jens Dick-Nielsen
Three Essays on Corporate Bond Market Liquidity

34. Sabrina Speiermann
Modstandens Politik Kampagnestyring i Velfærdsstaten. En diskussion af trafikkampagners styringspotentiale

35. Julie Uldam
Fickle Commitment. Fostering political engagement in 'the flighty world of online activism'

36. Annegrete Juul Nielsen
Traveling technologies and transformations in health care

37. Athur Mühlen-Schulte
Organising Development
Power and Organisational Reform in the United Nations Development Programme

38. Louise Rygaard Jonas
Branding på butiksgulvet
Et case-studie af kultur- og identitets-arbejdet i Kvickly

2011

1. Stefan Fraenkel
Key Success Factors for Sales Force Readiness during New Product Launch
A Study of Product Launches in the Swedish Pharmaceutical Industry

2. Christian Plesner Rossing
International Transfer Pricing in Theory and Practice

3. Tobias Dam Hede
Samtalekunst og ledelsesdisciplin
– en analyse af coachingsdiskursens genealogi og governmentality

4. Kim Pettersson
Essays on Audit Quality, Auditor Choice, and Equity Valuation

5. Henrik Merkelsen
The expert-lay controversy in risk research and management. Effects of institutional distances. Studies of risk definitions, perceptions, management and communication

6. Simon S. Torp
Employee Stock Ownership:
Effect on Strategic Management and Performance

7. Mie Harder
Internal Antecedents of Management Innovation

8. Ole Helby Petersen
Public-Private Partnerships: Policy and Regulation – With Comparative and Multi-level Case Studies from Denmark and Ireland

9. Morten Krogh Petersen
'Good' Outcomes. Handling Multiplicity in Government Communication

10. Kristian Tangsgaard Hvelplund
Allocation of cognitive resources in translation - an eye-tracking and key-logging study

11. Moshe Yonatany
The Internationalization Process of Digital Service Providers

12. Anne Vestergaard
Distance and Suffering
Humanitarian Discourse in the age of Mediatization

13. Thorsten Mikkelsen
Personligsheds indflydelse på forretningsrelationer

14. Jane Thostrup Jagd
Hvorfor fortsætter fusionsbølgen udover "the tipping point"?
– en empirisk analyse af information og kognitioner om fusioner

15. Gregory Gimpel
Value-driven Adoption and Consumption of Technology: Understanding Technology Decision Making

16. Thomas Stengade Sønderskov
Den nye mulighed
Social innovation i en forretningsmæssig kontekst

17. Jeppe Christoffersen
Donor supported strategic alliances in developing countries

18. Vibeke Vad Baunsgaard
Dominant Ideological Modes of Rationality: Cross functional

integration in the process of product innovation

19. Throstur Olaf Sigurjonsson
Governance Failure and Icelands's Financial Collapse

20. Allan Sall Tang Andersen
Essays on the modeling of risks in interest-rate and inflation markets

21. Heidi Tscherning
Mobile Devices in Social Contexts

22. Birgitte Gorm Hansen
Adapting in the Knowledge Economy
Lateral Strategies for Scientists and Those Who Study Them

23. Kristina Vaarst Andersen
Optimal Levels of Embeddedness
The Contingent Value of Networked Collaboration

24. Justine Grønbæk Pors
Noisy Management
A History of Danish School Governing from 1970-2010

25. Stefan Linder
Micro-foundations of Strategic Entrepreneurship
Essays on Autonomous Strategic Action

26. Xin Li
Toward an Integrative Framework of National Competitiveness
An application to China

27. Rune Thorbjørn Clausen
Værdifuld arkitektur
Et eksplorativt studie af bygningers rolle i virksomheders værdiskabelse

28. Monica Viken
Markedsundersøkelser som bevis i varemerke- og markedsføringsrett

29. Christian Wymann
Tattooing
The Economic and Artistic Constitution of a Social Phenomenon

30. Sanne Frandsen
Productive Incoherence
A Case Study of Branding and Identity Struggles in a Low-Prestige Organization

31. Mads Stenbo Nielsen
Essays on Correlation Modelling

32. Ivan Häuser
Følelse og sprog
Etablering af en ekspressiv kategori, eksemplificeret på russisk

33. Sebastian Schwenen
Security of Supply in Electricity Markets

2012

1. Peter Holm Andreasen
The Dynamics of Procurement Management
- A Complexity Approach

2. Martin Haulrich
Data-Driven Bitext Dependency Parsing and Alignment

3. Line Kirkegaard
Konsulenten i den anden nat
En undersøgelse af det intense arbejdsliv

4. Tonny Stenheim
Decision usefulness of goodwill under IFRS

5. Morten Lind Larsen
Produktivitet, vækst og velfærd
Industrirådet og efterkrigstidens Danmark 1945 - 1958

6. Petter Berg
Cartel Damages and Cost Asymmetries

7. Lynn Kahle
Experiential Discourse in Marketing
A methodical inquiry into practice and theory

8. Anne Roelsgaard Obling
Management of Emotions in Accelerated Medical Relationships

9. Thomas Frandsen
Managing Modularity of Service Processes Architecture

10. Carina Christine Skovmøller
CSR som noget særligt
Et casestudie om styring og meningsskabelse i relation til CSR ud fra en intern optik

11. Michael Tell
Fradragsbeskæring af selskabers finansieringsudgifter
En skatteretlig analyse af SEL §§ 11, 11B og 11C

12. Morten Holm
Customer Profitability Measurement Models
Their Merits and Sophistication across Contexts

13. Katja Joo Dyppel
Beskatning af derivater
En analyse af dansk skatteret

14. Esben Anton Schultz
Essays in Labor Economics
Evidence from Danish Micro Data

15. Carina Risvig Hansen
"Contracts not covered, or not fully covered, by the Public Sector Directive"

16. Anja Svejgaard Pors
Iværksættelse af kommunikation
- patientfigurer i hospitalets strategiske kommunikation

17. Frans Bévort
Making sense of management with logics
An ethnographic study of accountants who become managers

18. René Kallestrup
The Dynamics of Bank and Sovereign Credit Risk

19. Brett Crawford
Revisiting the Phenomenon of Interests in Organizational Institutionalism
The Case of U.S. Chambers of Commerce

20. Mario Daniele Amore
Essays on Empirical Corporate Finance

21. Arne Stjernholm Madsen
The evolution of innovation strategy
Studied in the context of medical device activities at the pharmaceutical company Novo Nordisk A/S in the period 1980-2008

22. Jacob Holm Hansen
Is Social Integration Necessary for Corporate Branding?
A study of corporate branding strategies at Novo Nordisk

23. Stuart Webber
Corporate Profit Shifting and the Multinational Enterprise

24. Helene Ratner
Promises of Reflexivity
Managing and Researching Inclusive Schools

25. Therese Strand
The Owners and the Power: Insights from Annual General Meetings

26. Robert Gavin Strand
In Praise of Corporate Social Responsibility Bureaucracy

27. Nina Sormunen
Auditor's going-concern reporting
Reporting decision and content of the report

28. John Bang Mathiasen
Learning within a product development working practice:
- an understanding anchored in pragmatism

29. Philip Holst Riis
Understanding Role-Oriented Enterprise Systems: From Vendors to Customers

30. Marie Lisa Dacanay
Social Enterprises and the Poor
Enhancing Social Entrepreneurship and Stakeholder Theory

31. Fumiko Kano Glückstad
Bridging Remote Cultures: Cross-lingual concept mapping based on the information receiver's prior-knowledge

32. Henrik Barslund Fosse
Empirical Essays in International Trade

33. Peter Alexander Albrecht
Foundational hybridity and its reproduction
Security sector reform in Sierra Leone

34. Maja Rosenstock
CSR - hvor svært kan det være?
Kulturanalytisk casestudie om udfordringer og dilemmaer med at forankre Coops CSR-strategi

35. Jeanette Rasmussen
Tweens, medier og forbrug
Et studie af 10-12 årige danske børns brug af internettet, opfattelse og forståelse af markedsføring og forbrug

36. Ib Tunby Gulbrandsen
'This page is not intended for a US Audience'
A five-act spectacle on online communication, collaboration & organization.

37. Kasper Aalling Teilmann
Interactive Approaches to Rural Development

38. Mette Mogensen
The Organization(s) of Well-being and Productivity
(Re)assembling work in the Danish Post

39. Søren Friis Møller
From Disinterestedness to Engagement
Towards Relational Leadership In the Cultural Sector

40. Nico Peter Berhausen
Management Control, Innovation and Strategic Objectives – Interactions and Convergence in Product Development Networks

41. Balder Onarheim
Creativity under Constraints
Creativity as Balancing 'Constrainedness'

42. Haoyong Zhou
Essays on Family Firms

43. Elisabeth Naima Mikkelsen
Making sense of organisational conflict
An empirical study of enacted sense-making in everyday conflict at work

2013

1. Jacob Lyngsie
Entrepreneurship in an Organizational Context

2. Signe Groth-Brodersen
Fra ledelse til selvet
En socialpsykologisk analyse af forholdet imellem selvledelse, ledelse og stress i det moderne arbejdsliv

3. Nis Høyrup Christensen
Shaping Markets: A Neoinstitutional Analysis of the Emerging Organizational Field of Renewable Energy in China

4. Christian Edelvold Berg
As a matter of size
THE IMPORTANCE OF CRITICAL MASS AND THE CONSEQUENCES OF SCARCITY FOR TELEVISION MARKETS

5. Christine D. Isakson
Coworker Influence and Labor Mobility
Essays on Turnover, Entrepreneurship and Location Choice in the Danish Maritime Industry

6. Niels Joseph Jerne Lennon
Accounting Qualities in Practice
Rhizomatic stories of representational faithfulness, decision making and control

7. Shannon O'Donnell
Making Ensemble Possible
How special groups organize for collaborative creativity in conditions of spatial variability and distance

8. Robert W. D. Veitch
Access Decisions in a Partly-Digital World
Comparing Digital Piracy and Legal Modes for Film and Music

9. Marie Mathiesen
Making Strategy Work
An Organizational Ethnography

10. Arisa Shollo
The role of business intelligence in organizational decision-making

11. Mia Kaspersen
The construction of social and environmental reporting

12. Marcus Møller Larsen
The organizational design of offshoring

13. Mette Ohm Rørdam
EU Law on Food Naming
The prohibition against misleading names in an internal market context

14. Hans Peter Rasmussen
GIV EN GED!
Kan giver-idealtyper forklare støtte til velgørenhed og understøtte relationsopbygning?

15. Ruben Schachtenhaufen
Fonetisk reduktion i dansk

16. Peter Koerver Schmidt
Dansk CFC-beskatning
I et internationalt og komparativt perspektiv

17. Morten Froholdt
Strategi i den offentlige sektor
En kortlægning af styringsmæssig kontekst, strategisk tilgang, samt anvendte redskaber og teknologier for udvalgte danske statslige styrelser

18. Annette Camilla Sjørup
Cognitive effort in metaphor translation
An eye-tracking and key-logging study

19. Tamara Stucchi
The Internationalization of Emerging Market Firms:
A Context-Specific Study

20. Thomas Lopdrup-Hjorth
"Let's Go Outside":
The Value of Co-Creation

21. Ana Alačovska
Genre and Autonomy in Cultural Production
The case of travel guidebook production

22. Marius Gudmand-Høyer
Stemningssindssygdommenes historie i det 19. århundrede
Omtydningen af melankolien og manien som bipolære stemningslidelser i dansk sammenhæng under hensyn til dannelsen af det moderne følelseslivs relative autonomi.
En problematiserings- og erfarings-analytisk undersøgelse

23. Lichen Alex Yu
Fabricating an S&OP Process
Circulating References and Matters of Concern

24. Esben Alfort
The Expression of a Need
Understanding search

25. Trine Pallesen
Assembling Markets for Wind Power
An Inquiry into the Making of Market Devices

26. Anders Koed Madsen
Web-Visions
Repurposing digital traces to organize social attention

27. Lærke Højgaard Christiansen
BREWING ORGANIZATIONAL RESPONSES TO INSTITUTIONAL LOGICS

28. Tommy Kjær Lassen
EGENTLIG SELVLEDELSE
En ledelsesfilosofisk afhandling om selvledelsens paradoksale dynamik og eksistentielle engagement

29. Morten Rossing
Local Adaption and Meaning Creation in Performance Appraisal

30. Søren Obed Madsen
Lederen som oversætter
Et oversættelsesteoretisk perspektiv på strategisk arbejde

31. Thomas Høgenhaven
Open Government Communities
Does Design Affect Participation?

32. Kirstine Zinck Pedersen
Failsafe Organizing?
A Pragmatic Stance on Patient Safety

33. Anne Petersen
Hverdagslogikker i psykiatrisk arbejde
En institutionsetnografisk undersøgelse af hverdagen i psykiatriske organisationer

34. Didde Maria Humle
Fortællinger om arbejde

35. Mark Holst-Mikkelsen
Strategieksekvering i praksis
– barrierer og muligheder!

36. Malek Maalouf
Sustaining lean
Strategies for dealing with organizational paradoxes

37. Nicolaj Tofte Brenneche
Systemic Innovation In The Making
The Social Productivity of Cartographic Crisis and Transitions in the Case of SEEIT

38. Morten Gylling
The Structure of Discourse
A Corpus-Based Cross-Linguistic Study

39. Binzhang YANG
Urban Green Spaces for Quality Life
- Case Study: the landscape architecture for people in Copenhagen

40. Michael Friis Pedersen
Finance and Organization:
The Implications for Whole Farm Risk Management

41. Even Fallan
Issues on supply and demand for environmental accounting information

42. Ather Nawaz
Website user experience
A cross-cultural study of the relation between users´ cognitive style, context of use, and information architecture of local websites

43. Karin Beukel
The Determinants for Creating Valuable Inventions

44. Arjan Markus
External Knowledge Sourcing and Firm Innovation
Essays on the Micro-Foundations of Firms' Search for Innovation

2014

1. Solon Moreira
Four Essays on Technology Licensing and Firm Innovation

2. Karin Strzeletz Ivertsen
Partnership Drift in Innovation Processes
A study of the Think City electric car development

3. Kathrine Hoffmann Pii
Responsibility Flows in Patient-centred Prevention

4. Jane Bjørn Vedel
Managing Strategic Research
An empirical analysis of science-industry collaboration in a pharmaceutical company

5. Martin Gylling
Processuel strategi i organisationer
Monografi om dobbeltheden i tænkning af strategi, dels som vidensfelt i organisationsteori, dels som kunstnerisk tilgang til at skabe i erhvervsmæssig innovation

6. Linne Marie Lauesen
Corporate Social Responsibility in the Water Sector: How Material Practices and their Symbolic and Physical Meanings Form a Colonising Logic

7. Maggie Qiuzhu Mei
LEARNING TO INNOVATE: The role of ambidexterity, standard, and decision process

8. Inger Høedt-Rasmussen
Developing Identity for Lawyers Towards Sustainable Lawyering

9. Sebastian Fux
Essays on Return Predictability and Term Structure Modelling

10. Thorbjørn N. M. Lund-Poulsen
Essays on Value Based Management

11. Oana Brindusa Albu
Transparency in Organizing: A Performative Approach

12. Lena Olaison
Entrepreneurship at the limits

13. Hanne Sørum
DRESSED FOR WEB SUCCESS? An Empirical Study of Website Quality in the Public Sector

14. Lasse Folke Henriksen
Knowing networks How experts shape transnational governance

15. Maria Halbinger
Entrepreneurial Individuals Empirical Investigations into Entrepreneurial Activities of Hackers and Makers

16. Robert Spliid
Kapitalfondenes metoder og kompetencer

17. Christiane Stelling
Public-private partnerships & the need, development and management of trusting A processual and embedded exploration

18. Marta Gasparin
Management of design as a translation process

19. Kåre Moberg
Assessing the Impact of Entrepreneurship Education From ABC to PhD

20. Alexander Cole
Distant neighbors Collective learning beyond the cluster

21. Martin Møller Boje Rasmussen
Is Competitiveness a Question of Being Alike? How the United Kingdom, Germany and Denmark Came to Compete through their Knowledge Regimes from 1993 to 2007

22. Anders Ravn Sørensen
Studies in central bank legitimacy, currency and national identity Four cases from Danish monetary history

23. Nina Bellak
Can Language be Managed in International Business? Insights into Language Choice from a Case Study of Danish and Austrian Multinational Corporations (MNCs)

24. Rikke Kristine Nielsen
Global Mindset as Managerial Meta-competence and Organizational Capability: Boundary-crossing Leadership Cooperation in the MNC The Case of 'Group Mindset' in Solar A/S.

25. Rasmus Koss Hartmann
User Innovation inside government Towards a critically performative foundation for inquiry

26. Kristian Gylling Olesen
Flertydig og emergerende ledelse i folkeskolen
Et aktør-netværksteoretisk ledelses-studie af politiske evalueringsreformers betydning for ledelse i den danske folkeskole

27. Troels Riis Larsen
Kampen om Danmarks omdømme 1945-2010
Omdømmearbejde og omdømmepolitik

28. Klaus Majgaard
Jagten på autenticitet i offentlig styring

29. Ming Hua Li
Institutional Transition and Organizational Diversity: Differentiated internationalization strategies of emerging market state-owned enterprises

30. Sofie Blinkenberg Federspiel
IT, organisation og digitalisering: Institutionelt arbejde i den kommunale digitaliseringsproces

31. Elvi Weinreich
Hvilke offentlige ledere er der brug for når velfærdstænkningen flytter sig – er Diplomuddannelsens lederprofil svaret?

32. Ellen Mølgaard Korsager
Self-conception and image of context in the growth of the firm – A Penrosian History of Fiberline Composites

33. Else Skjold
The Daily Selection

34. Marie Louise Conradsen
The Cancer Centre That Never Was
The Organisation of Danish Cancer Research 1949-1992

35. Virgilio Failla
Three Essays on the Dynamics of Entrepreneurs in the Labor Market

36. Nicky Nedergaard
Brand-Based Innovation
Relational Perspectives on Brand Logics and Design Innovation Strategies and Implementation

37. Mads Gjedsted Nielsen
Essays in Real Estate Finance

38. Kristin Martina Brandl
Process Perspectives on Service Offshoring

39. Mia Rosa Koss Hartmann
In the gray zone
With police in making space for creativity

40. Karen Ingerslev
Healthcare Innovation under The Microscope
Framing Boundaries of Wicked Problems

41. Tim Neerup Themsen
Risk Management in large Danish public capital investment programmes

2015

1. Jakob Ion Wille
Film som design
Design af levende billeder i film og tv-serier

2. Christiane Mossin
Interzones of Law and Metaphysics
Hierarchies, Logics and Foundations of Social Order seen through the Prism of EU Social Rights

3. Thomas Tøth
TRUSTWORTHINESS: ENABLING GLOBAL COLLABORATION
An Ethnographic Study of Trust, Distance, Control, Culture and Boundary Spanning within Offshore Outsourcing of IT Services

4. Steven Højlund
Evaluation Use in Evaluation Systems – The Case of the European Commission

5. Julia Kirch Kirkegaard
AMBIGUOUS WINDS OF CHANGE – OR FIGHTING AGAINST WINDMILLS IN CHINESE WIND POWER
A CONSTRUCTIVIST INQUIRY INTO CHINA'S PRAGMATICS OF GREEN MARKETISATION MAPPING CONTROVERSIES OVER A POTENTIAL TURN TO QUALITY IN CHINESE WIND POWER

6. Michelle Carol Antero
A Multi-case Analysis of the Development of Enterprise Resource Planning Systems (ERP) Business Practices

Morten Friis-Olivarius
The Associative Nature of Creativity

7. Mathew Abraham
New Cooperativism:
A study of emerging producer organisations in India

8. Stine Hedegaard
Sustainability-Focused Identity: Identity work performed to manage, negotiate and resolve barriers and tensions that arise in the process of constructing or ganizational identity in a sustainability context

9. Cecilie Glerup
Organizing Science in Society – the conduct and justification of resposible research

10. Allan Salling Pedersen
Implementering af ITIL® IT-governance - når best practice konflikter med kulturen Løsning af implementerings- problemer gennem anvendelse af kendte CSF i et aktionsforskningsforløb.

11. Nihat Misir
A Real Options Approach to Determining Power Prices

12. Mamdouh Medhat
MEASURING AND PRICING THE RISK OF CORPORATE FAILURES

13. Rina Hansen
Toward a Digital Strategy for Omnichannel Retailing

14. Eva Pallesen
In the rhythm of welfare creation
A relational processual investigation moving beyond the conceptual horizon of welfare management

15. Gouya Harirchi
In Search of Opportunities: Three Essays on Global Linkages for Innovation

16. Lotte Holck
Embedded Diversity: A critical ethnographic study of the structural tensions of organizing diversity

17. Jose Daniel Balarezo
Learning through Scenario Planning

18. Louise Pram Nielsen
Knowledge dissemination based on terminological ontologies. Using eye tracking to further user interface design.

TITLER I ATV PH.D.-SERIEN

1992
1. Niels Kornum
Servicesamkørsel – organisation, økonomi og planlægningsmetode

1995
2. Verner Worm
Nordiske virksomheder i Kina
Kulturspecifikke interaktionsrelationer ved nordiske virksomhedsetableringer i Kina

1999
3. Mogens Bjerre
Key Account Management of Complex Strategic Relationships
An Empirical Study of the Fast Moving Consumer Goods Industry

2000
4. Lotte Darsø
Innovation in the Making
Interaction Research with heterogeneous Groups of Knowledge Workers creating new Knowledge and new Leads

2001
5. Peter Hobolt Jensen
Managing Strategic Design Identities
The case of the Lego Developer Network

2002
6. Peter Lohmann
The Deleuzian Other of Organizational Change – Moving Perspectives of the Human

7. Anne Marie Jess Hansen
To lead from a distance: The dynamic interplay between strategy and strategizing – A case study of the strategic management process

2003
8. Lotte Henriksen
Videndeling
– om organisatoriske og ledelsesmæssige udfordringer ved videndeling i praksis

9. Niels Christian Nickelsen
Arrangements of Knowing: Coordinating Procedures Tools and Bodies in Industrial Production – a case study of the collective making of new products

2005
10. Carsten Ørts Hansen
Konstruktion af ledelsesteknologier og effektivitet

TITLER I DBA PH.D.-SERIEN

2007
1. Peter Kastrup-Misir
Endeavoring to Understand Market Orientation – and the concomitant co-mutation of the researched, the re searcher, the research itself and the truth

2009
1. Torkild Leo Thellefsen
Fundamental Signs and Significance effects
A Semeiotic outline of Fundamental Signs, Significance-effects, Knowledge Profiling and their use in Knowledge Organization and Branding

2. Daniel Ronzani
When Bits Learn to Walk Don't Make Them Trip. Technological Innovation and the Role of Regulation by Law in Information Systems Research: the Case of Radio Frequency Identification (RFID)

2010
1. Alexander Carnera
Magten over livet og livet som magt
Studier i den biopolitiske ambivalens