Maps, how do users see them?

An in depth investigation of the map users' cognitive processes

Kristien Ooms
Maps, how do users see them?

An in depth investigation of the map users’ cognitive processes

Kristien Ooms
Copyright © Kristien Ooms, Department of Geography, Faculty of Sciences, Ghent University 2012. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without permission in writing from the copyright holder(s).

The research reported in this dissertation was conducted at the CartoGIS research unit, Department of Geography, Faculty of Sciences, Ghent University.

Map on cover – modified from:
NGI Belgium
Map sheet 49/2S (Comblain-au-Pont)
Original scale 1 : 10 000
Authorization number A2706
Maps, how do users see them?
An in depth investigation of the map users’
cognitive processes

Dissertation submitted in accordance with the requirements for the
degree of Doctor of Sciences: Geomatics and Surveying

Kaarten, hoe zien gebruikers ze?
Een diepgaand onderzoek naar de cognitieve processen
van kaartgebruikers

Proefschrift aangeboden tot het behalen van de graad van Doctor in de
Wetenschappen: Geomatica en Landmeetkunde

Kristien Ooms
Supervisors
Prof. dr. Philippe De Maeyer  
*Department of Geography, Ghent University*

Prof. dr. Veerle Fack  
*Department of Applied Mathematics and Computer Science, Ghent University*

Members of the Reading Committee
Prof. dr. Sara Fabrikant  
*GIVA, Department of Geography, University of Zurich*

Prof. dr. Georg Gartner  
*President ICA, Institute of Geoinformation and Cartography, Vienna University of Technology*

Prof. dr. Veerle Van Eetvelde  
*Department of Geography, Ghent University*

Members of the Examination Committee
Prof. dr. ir. Alain De Wulf (Chairman)  
*Department of Geography, Ghent University*

Prof. dr. Nico Van de Weghe  
*Department of Geography, Ghent University*

Prof. dr. Frank Witlox  
*Department of Geography, Ghent University*
# Table of Contents

List of Figures ........................................................................................................................................ v
List of Tables ......................................................................................................................................... ix
Preface .................................................................................................................................................. xi

Chapter 1: Introduction .......................................................................................................................... 1
  1.1. Background information .............................................................................................................. 2
      1.1.1. The communication of cartographic information .......................................................... 2
      1.1.2. Cartography and cognition ............................................................................................... 6
      1.1.3. Improving the effectiveness of maps: UCD ...................................................................... 8
  1.2. Rationale and synopsis ................................................................................................................ 11
      1.2.1. Research objectives and questions .................................................................................. 11
      1.2.2. Dissertation outline ....................................................................................................... 14
      1.2.3. Out of scope .................................................................................................................. 17

References .............................................................................................................................................. 19

Chapter 2: Interpreting Maps Through the Eyes of Expert and Novice Users ...................................... 25
  2.1. Introduction .................................................................................................................................. 26
  2.2. Study design .................................................................................................................................. 30
      2.2.1. Participants ....................................................................................................................... 30
      2.2.2. Stimuli ............................................................................................................................... 30
      2.2.3. Procedure .......................................................................................................................... 33
      2.2.4. Apparatus .......................................................................................................................... 34
      2.2.5. Recordings ......................................................................................................................... 34
  2.3. Results .......................................................................................................................................... 34
      2.3.1. Response time measurements ......................................................................................... 34
      2.3.2. Fixation duration .............................................................................................................. 37
      2.3.3. Fixation count ................................................................................................................... 38
      2.3.4. Fixation distribution ........................................................................................................ 40
  2.4. Discussion ...................................................................................................................................... 41
  2.5. Conclusion and future work ........................................................................................................ 44

References .............................................................................................................................................. 45

Chapter 3: Analysing the Spatial Dimension of Eye Movement Data Using a Visual Analytic Approach .............................................................................................................................................................................. 49
  3.1. Introduction .................................................................................................................................. 50

<table>
<thead>
<tr>
<th>Chapter 4: Investigating the Effectiveness of an Efficient Label Placement Method Using Eye Movement Data</th>
<th>65</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1. Introduction ........................................................................................................ 66</td>
<td></td>
</tr>
<tr>
<td>4.1.1. Towards efficient cartographic displays ........................................ 67</td>
<td></td>
</tr>
<tr>
<td>4.1.2. Improved efficiency .............................................................................. 71</td>
<td></td>
</tr>
<tr>
<td>4.1.3. Tracking effectiveness .......................................................................... 75</td>
<td></td>
</tr>
<tr>
<td>4.2. Study design ......................................................................................................... 76</td>
<td></td>
</tr>
<tr>
<td>4.2.1. Participants ................................................................................................. 77</td>
<td></td>
</tr>
<tr>
<td>4.2.2. Task and stimuli ........................................................................................... 77</td>
<td></td>
</tr>
<tr>
<td>4.2.3. Apparatus and data ...................................................................................... 81</td>
<td></td>
</tr>
<tr>
<td>4.3. Results and discussion ....................................................................................... 83</td>
<td></td>
</tr>
<tr>
<td>4.3.1. Relative response times ............................................................................... 84</td>
<td></td>
</tr>
<tr>
<td>4.3.2. Fixation count and duration ....................................................................... 85</td>
<td></td>
</tr>
<tr>
<td>4.3.3. Qualitative scanpath analysis .................................................................... 87</td>
<td></td>
</tr>
<tr>
<td>4.3.4. Questionnaire .............................................................................................. 89</td>
<td></td>
</tr>
<tr>
<td>4.4. Conclusion and future work ............................................................................... 89</td>
<td></td>
</tr>
</tbody>
</table>

Chapter 5: Understanding Expert and Novice Map Users: Reading and Interpretation |
---------------------------------------------------------------|-----|
<p>| 5.1. Introduction ......................................................................................................... 98 |
| 5.1.1. How can we process visual stimuli? ............................................................ 99 |
| 5.1.2. How do we interpret maps? ........................................................................ 101 |
| 5.1.3. How to study ‘map interpretation’? ............................................................ 104 |
| 5.2. Study design ......................................................................................................... 105 |
| 5.2.1. Participants .................................................................................................. 105 |
| 5.2.2. Tasks .......................................................................................................... 106 |
| 5.2.3. Stimuli ......................................................................................................... 107 |</p>
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.4. Apparatus and recordings</td>
<td>108</td>
</tr>
<tr>
<td>5.3. Methodology and results</td>
<td>109</td>
</tr>
<tr>
<td>5.3.1. Statistical comparison: experts vs. novices</td>
<td>109</td>
</tr>
<tr>
<td>5.3.2. Heatmaps</td>
<td>110</td>
</tr>
<tr>
<td>5.3.3. Gridded visualisation: methodology</td>
<td>111</td>
</tr>
<tr>
<td>5.3.4. Gridded visualisation: total fixation count</td>
<td>114</td>
</tr>
<tr>
<td>5.3.5. Gridded visualisation: total dwell time</td>
<td>116</td>
</tr>
<tr>
<td>5.3.6. 3D gridded visualisation: total dwell time</td>
<td>117</td>
</tr>
<tr>
<td>5.3.7. 3D gridded visualisation: fixation duration per fixation</td>
<td>120</td>
</tr>
<tr>
<td>5.3.8. Statistical grids: mirrored maps</td>
<td>122</td>
</tr>
<tr>
<td>5.3.9. Scanpath visualisation</td>
<td>125</td>
</tr>
<tr>
<td>5.4. Discussion and conclusion</td>
<td>128</td>
</tr>
<tr>
<td>References</td>
<td>132</td>
</tr>
<tr>
<td>Chapter 6: Listen to the Map User: Cognition, Memory, and Expertise</td>
<td>135</td>
</tr>
<tr>
<td>6.1. Introduction</td>
<td>136</td>
</tr>
<tr>
<td>6.1.1. Understanding the map users' cognitive structures</td>
<td>137</td>
</tr>
<tr>
<td>6.1.2. Influence of expertise</td>
<td>138</td>
</tr>
<tr>
<td>6.1.3. Accessing the users' cognitive processes</td>
<td>140</td>
</tr>
<tr>
<td>6.1.4. Research objective and question</td>
<td>143</td>
</tr>
<tr>
<td>6.2. Materials and methods</td>
<td>144</td>
</tr>
<tr>
<td>6.2.1. Participants</td>
<td>144</td>
</tr>
<tr>
<td>6.2.2. Tasks and stimuli</td>
<td>144</td>
</tr>
<tr>
<td>6.2.3. Apparatus and recordings</td>
<td>147</td>
</tr>
<tr>
<td>6.2.4. Methodology</td>
<td>147</td>
</tr>
<tr>
<td>6.3. Results</td>
<td>155</td>
</tr>
<tr>
<td>6.3.1. Word segmentation</td>
<td>155</td>
</tr>
<tr>
<td>6.3.2. Sentence segmentation: coded protocols</td>
<td>159</td>
</tr>
<tr>
<td>6.3.3. Sentence segmentation: contexts</td>
<td>162</td>
</tr>
<tr>
<td>6.3.4. Score on drawn maps</td>
<td>163</td>
</tr>
<tr>
<td>6.3.5. Order of drawing</td>
<td>165</td>
</tr>
<tr>
<td>6.3.6. Questionnaire results</td>
<td>167</td>
</tr>
<tr>
<td>6.4. Discussion</td>
<td>168</td>
</tr>
<tr>
<td>6.5. Conclusion and future work</td>
<td>171</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>172</td>
</tr>
</tbody>
</table>
References .................................................................................................................................................... 172

Chapter 7: General Discussion ..................................................................................................................... 177

7.1. Summary ........................................................................................................................................... 178
7.2. Improving the effectiveness of (screen) maps ....................................................................................... 187
7.3. Critical reflection on the methodology .................................................................................................. 189
7.4. Recommendations for future work ........................................................................................................ 191
   7.4.1. Future work considering alternative tasks .................................................................................... 191
   7.4.2. Future work considering map users ........................................................................................... 192
   7.4.3. Future work considering map types .......................................................................................... 192
   7.4.4. Future work considering user interactions ................................................................................. 193
References .................................................................................................................................................... 194

Chapter 8: General Conclusion ..................................................................................................................... 199

Samenvatting (Dutch Summary) .................................................................................................................. 205

Referenties ................................................................................................................................................... 211

Curriculum Vitae .......................................................................................................................................... 213
List of Figures

Figure 1.1: Model of cartographic communication (after Kolácný, 1969) ......................... 2
Figure 1.2: A typical UCD-lifecycle .................................................................................. 9
Figure 1.3: Dissertation outline ...................................................................................... 15
Figure 2.1: Initial (a) and final (b) view of a demo-map ................................................... 31
Figure 2.2: Mean values (with 95% confidence interval) of the relative response times for locating the subsequent labels ................................................................. 35
Figure 2.3: Mean values (with 95% confidence interval) of the fixation durations related to each subsequent time interval ................................................................. 38
Figure 2.4: Mean values (with 95% confidence interval) of the fixation counts (fix/s) related to each subsequent time interval ................................................................. 39
Figure 2.5: Distribution of the fixations of both user groups on the map image (in percentage) during a fixed time interval ................................................................. 41
Figure 3.1: Example of a demo-map with on top the initial view, bottom left a view during the pan-operation and bottom right the final view .............................................. 53
Figure 3.2: Overcrowded visualisation if all eye movements are selected ...................... 54
Figure 3.3: Overcrowded visualisation when minor selections are applied ...................... 55
Figure 3.4: Time series to visualise the evolution of the scanpaths over time (in s) .... 56
Figure 3.5: Filtering the scanpaths based on two attributes: participant and demo-map .................................................................................................................. 57
Figure 3.6: Filtering the scanpaths based on one demo-map and time interval .......... 58
Figure 3.7: Generalisation and summarisation of the scanpaths ...................................... 59
Figure 3.8: Simplification of the scanpaths based on crossed areas ................................ 60
Figure 3.9: Aggregation of the scanpaths based on crossed areas ................................. 60
Figure 3.10: Division of the scanpaths to extract individual trajectories between visited areas ........................................................................................................... 62
Figure 4.1: Illustration of the number of labels considered by the algorithms during panning .................................................................................................................. 72
Figure 4.2: Illustration of the three subsequent intervals of the map stimulus .......... 78
Figure 4.3: Overview of the users’ cognitive processes during the visual search task .......................................................................................................................... 80
List of Figures

Figure 4.4: Mean response time measurements (with 95% confidence interval) for both types of label placement related to finding the five labels, before (a) and after (b) the interaction.........................................................84
Figure 4.5: Mean fixation count (fix/s, with 95% confidence interval) for both types of label placement, before and after the interaction.........................................................86
Figure 4.6: Mean fixation duration (with 95% confidence interval) for both types of label placement, before and after the interaction ............................................................87
Figure 4.7: The users’ scanpaths when searching on one of the maps ........................................88
Figure 5.1: Memory structure: an integrated model according to Atkinson and Shiffrin (1968) and Baddely (1999) ..................................................................................................................101
Figure 5.2: Stimuli depicted during the trials of the user study................................................108
Figure 5.3: Heatmaps from the same participants for the four trials..................................110
Figure 5.4: Grid of AOIs placed over each map image (yellow) to visually analyse the users’ fixations (cyan) ...............................................................112
Figure 5.5: Grid of AOIs depicting the average fixation count for experts (top row) and novices (bottom row) for each stimuli (left to right)..............................115
Figure 5.6: Grid of AOIs depicting the maximum fixation count for experts (top row) and novices (bottom row) for each stimuli (left to right) ......................115
Figure 5.7: Grid of AOIs depicting the average total dwell time (map 2)...............................117
Figure 5.8: 3D representation of the average total dwell time for map 1 and map 4. 118
Figure 5.9: 3D representation of the average total dwell time for map 2 and map 3. 118
Figure 5.10: 3D representation of the average fixation duration per fixation for the expert users ....................................................................................................121
Figure 5.11: 3D representation of the average fixation duration per fixation for the novice users ........................................................................................................122
Figure 5.12: Statistical comparison of the number of fixation between two mirrored maps.................................................................................................124
Figure 5.13: Evolution of the participants’ scanpaths for the different maps and user groups ................................................................................................................124
Figure 6.1: Analysing verbal protocols obtained from a thinking aloud study .................142
Figure 6.2: The stimuli presented on screen during the initial part of the trial ...............146
Figure 6.3: Highest (general) level of the psychological model ........................................149
Figure 6.4: Psychological model of the assignment execution ........................................150
Figure 6.5: Psychological model of the evaluation process ............................................150
Figure 6.6: Psychological model of the information retrieval process ....................... 151
Figure 6.7: Overview of the ordered word segmentation (first 50% of words) and their differences in word frequencies between expert and novice users . 157
Figure 6.8: Overview of the word segmentation, based on a thematic categorisation ............................................................................................... 158
Figure 6.9: Overview of the segmented protocols (full thought) per level and map 160
Figure 6.10: Overview of the scores (%) on the drawn maps, per theme (left) and per map (right) ........................................................................................................ 164
Figure 6.11: The order with which the objects were drawn for each map .............. 166
Figure 6.12: Result of the questionnaire regarding the confidence of the result ...... 168
List of Tables

Table 2.1: Mean values for the time intervals (in s) for locating subsequent labels and their statistical comparison.................................................................36
Table 2.2: Mean values for the fixation durations (in s) for locating subsequent labels, both before and after the interaction................................................38
Table 2.3: Mean values for the fixation counts (fix/s) for locating subsequent labels, both before and after the interaction.................................................39
Table 5.1: Statistical comparison (average fixation duration; number of fixations per second; duration of the trial) between expert and novice map users 110
Table 6.1: Proposed codes related to the map level (Level 1) ........................................152
Table 6.2: Proposed codes related to the item level (Level 2) .....................................152
Table 6.3: Proposed codes related to the participants’ level of confidence (Level 3) 153
Table 6.4: Proposed codes related to the participants’ actions (Level 4) .................153
Table 6.5: Proposed codes related to the word segmentation (Level 5) ...............153
Table 7.1: Overview of the dissertations’ content......................................................179
Preface

This first section in my dissertation is actually the last one that I am writing. It also feels like the end of an important chapter in my life. These past five years have been full of ups and downs, challenges and opportunities, stressful and wonderful moments. I learned so much during this period regarding the subject of my PhD, but also regarding the world around me. This knowledge is a priceless reward and I think that pursuing this PhD has also changed me as a person. Along the way I have met many interesting people, who have inspired me, supported me, guided me, and provided me with the much needed distractions. It would be impossible to name them all here, but I would like to mention a handful of them who contributed in some way to this work and thus shaped this dissertation.

One person I would like to thank in particularly said about six years ago that ‘I should start thinking about my future’ and that a PhD would be an option. At that time, I did not believe that I could accomplish something like that, but that future is now and it turns out that he was right. So, many thanks go to my supervisor, Philippe De Maeyer, for believing in me and giving me this opportunity. I am also indebted to my supervisor Veerle Fack, who guided me from a non-cartographic point of view. This interdisciplinary aspect has proven to be a key piece in the success of this PhD. During the day to day life in the office, I got to know both of my supervisors as very caring, supporting, and warm persons. Completing the dissertation would not have been possible without this.

I am also very grateful that I have so many open minded and enthusiastic colleagues, who create a nice atmosphere to work in and who happily participated in my user studies. It is not possible to name them all, but some of them deserve a special treatment because of their huge support during the past years. In particularly, I would like to thank Ann, Bart, Els, Ruben, and Wim for their aids in the educational matters. Furthermore, I would like to thank the persons with whom I spend most time in the office and who, consequently, helped to shape this dissertation in many ways: Leen, Ilke, Annelies, Lieselot, and Ellen.
A PhD is closely linked with quite a number of administrative issues and a pile of paper work. Thanks to the competence of the department’s secretary, Helga Vermeulen, I did not have to worry much about this as it always turns out right. Next to the administrative tasks, she is also always a listening ear where I can stop by for a small chat or when I need to get something of my chest. Thanks Helga!

I would like to express my gratitude to a number of people who made this PhD possible and supported me along the way. First I would like to thank the people of the Department of Experimental Psychology (especially Eva and Lise) for giving me the opportunity to work with their eye tracking device and guiding me in this. I am also indebted to the National Geographical Institute who provided me with the necessary data for my user studies and to Frédérique Spitaels in particularly. It is really nice to keep in touch and exchange thoughts with someone who works in the ‘real’ and ‘practical’ side of the cartographic world. This had a major influence on how the topic of my PhD ‘evolved’ over time.

During the last year and a half of my PhD, I also had the privilege to work together with Corné van Elzakker, Alex Pucher, and David Forrest in the ICA Commission on Use and User Issues. I must say that this has created a number of wonderful opportunities for me, which I otherwise would not have experienced. What is more, it is always nice working with you guys and I hope we can continue doing this in the future.

This PhD would also not have been possible without a little help from my friends. Beside my working hours, they provided me with the much needed distractions and I would like to thank them for all the fantastic times we had together during the past five years and in particularly my old ‘lmk’ study mates, Kim, Caroline, Stephanie, Koen, Evy, and Steven.

Special gratitude goes to my family – my parents, Miet, Marc, Joris, Brenda – who gave me the opportunity to study Geomatics on the other side of the country. They have always supported me in everything I did, even when I decided to stay in Ghent after my studies. I have been living here for about ten years, but I know now that I will always stay a ‘Limburger’, wherever I might end up in the future. I would also
like to mention Dirk and Erna here, and thank them for their support during the past few years.

One of the most important persons I met along the way is my boyfriend, Sven Schelfaut. People say that pursuing a PhD is a lonely task, but you showed me that this should not be the case. Meeting you was the best thing that happened to me. I want to thank you for your never ending love, support and, patience. You made sure that I could put my work aside at home and you inspired me with your contagious liveliness. The end of this PhD feels like closing a chapter of my life, but this means that we can begin a new one together…in Wachtebeke.

Ghent, November 26, 2012
Chapter 1

Introduction
1.1. Background information

1.1.1. The communication of cartographic information

Maps have already been used for dozens of centuries to (visually) present information with a spatial context. These cartographic products have evolved drastically over time: the medium on which they are presented, their design, their application fields, possibilities regarding animations and interactivity, etc. However, their main purpose has remained the same: communication of geographic information. The information around us has to be generalised and symbolised in order to present its spatial context on a medium with limited dimensions. However, this generalisation and symbolisation process might introduce ‘noise’ in the communication process (Dodge, et al., 2011; Keates, 1996; Muehrcke, 1986; among others).

In his book *The Look of Maps*, Robinson (1952) was one of the first to study map design in the light of the communication process. He recognised that the design of a map can have a strong impact on how the information is perceived and interpreted, and thus on the quality of the communication process. His work had a strong influence on cartographic research, even up to now (Montello, 2002). Based on this idea, an influential communication model for cartographic information was proposed by Kolácný (1969). A simplified version of this model is depicted in Figure 1.1. For a long time, his model was considered an essential aid in understanding how the cartographic information is transferred to the map user.

![Figure 1.1: Model of cartographic communication (after Kolácný, 1969)](image_url)
Kolácný (1969) depicted in his model the cartographer who observes only a small section of the reality, which he tries to visualise in a map using the cartographic language. The map user uses his knowledge of the cartographic language to decipher the map and recreate the depicted reality. The map creation and interpretation process is, however, influenced by a number of factors, such as knowledge, experience, psychological processes, abilities, aims, interests, tasks, and other external conditions. Due to these influencing factors, the recreated reality of the map user will not correspond to the (original) reality observed by the cartographer, causing errors in the communication process.

In 1967, Jacques Bertin proposed a number of strict cartographic rules that should be followed in order to effectively communicate the cartographic information. Bertin identified six graphical variables (size, shape, value, hue, orientation, and texture) that can be used to present an object’s property. He linked these to a number of visual properties related to their appropriate use (selection, association, order, and quantity). The application of different graphical variables creates different meanings in the map users’ minds: graphical variables have particular perception properties and these are linked to the characteristics (meanings) of the information to be represented. These ‘rules’ are still followed during the map making process of modern cartographic products. However, these rules are based on theoretical concepts, not on in-depth user research. Whereas Bertin focussed his research on the design of map symbols, other authors investigated cartographic design issues related to, among others, relief visualisations and the placement of name labels in the map image (Cuenin, 1972; Imhof, 1975; among others).

In the course of the cartographic development, it became clear that the context in which a cartographic product is to be used should not be ignored during the design process. The map’s content and its design should be keyed to, among others, the medium on which the map is presented, the type of information (thematic or topographic), and the target audience (Robinson, et al., 1995; Slocum, et al., 2005). In the past, almost all maps were printed on paper (atlases, posters, schoolbooks, etc.). Modern cartographic products can be found on a wide range of media, including computer screens (smartphones, tablets, notebooks, desktop PCs, touch
During the last few decades, the maps found on the Internet have greatly outnumbered their paper counterparts (Peterson, 2003).

These digital (screen based) maps enable new cartographic possibilities, such as animations and user interactions. Around the turn of the century, a number of books emerged that extensively covered and explored (potential) possibilities of ‘modern maps’: *Interactive and Animated Cartography* (Peterson, 1995), *Multimedia Cartography* (Cartwright, *et al.*, 1999), *Web Cartography* (Kraak & Brown, 2001), *Maps and the Internet* (Peterson, 2003), etc. Kraak (2001) identified a rough classification for the cartographic products found on the Internet, distinguishing between static and dynamic maps, on the one hand, and view-only and interactive maps, on the other hand. At that time he stated that the ‘static-view only map’ was most commonly found on the Internet. Ten years later – as a result of the fast evolving technical developments and the users’ needs to visualise the temporal component of geographic data – most maps found on the Internet are interactive and often dynamic.

In spite of the success of Internet cartography, screen maps come with a number of serious drawbacks and limitations as opposed to the paper medium: smaller dimensions of the screens, lower resolutions, and different colour definitions, to name a few (Cartwright, 1999; Peterson, 1995; among others).

Over the years, a number of critical reflections on Bertin’s original graphical variables appeared and additions and refinements were proposed. Excellent overviews of these reflections can be found in MacEachren (1995), Koch (2001), Kraak and Ormeling (2010), among others. Morrison (1974), for example, proposed to add saturation (besides hue and value) as a graphical variable. Recently, Bertin’s graphical variables were also extended in the light of the modern, digital cartography. Both the limitations (e.g. low output resolutions) and new possibilities are taken into account here, such as the dynamic variation of a variable over time: movement, enlargement, sound, etc. (Peterson, 1995).

Today, maps are distributed on a worldwide digital network, using the Internet. This makes them more accessible in comparison to paper maps. Analogue, printed maps
are ‘real’ objects that have to be moved physically from one location to another, which makes them less accessible. Digital products, such as web maps, can travel around the world with a single mouse click. Furthermore, one map can be consulted by thousands of users at the same time. This makes them more accessible to all kinds of users; from experts to novices. What is more, it is far easier to keep digital information up to date (van Elzakker, 2004). This increased accessibility and actuality also encourages the use of maps in a wider range of applications: navigation, policy, nature preservation, etc.

Even more recently, the term neocartography has appeared in research literature. Neocartography is closely linked with Web 2.0, which introduced a considerable change in how we interact with the information on the Internet. In the past, this was often a static one-way communication process that allowed the users to retrieve information. With Web 2.0, this process has become much more dynamic and interactive. The Internet users themselves can upload and share information, which can, in turn trigger reactions from other users. This also holds true in the new cartography of Web 2.0. New tools are available (such as APIs, Web Feature Services, Web Mapping Services), giving basically all web users the opportunity to create web maps (such as pushpin maps or mashup maps) or contribute to the content of web maps (such as Open Street Maps) through crowd sourcing. Nevertheless, most of these new ‘map makers’ did not receive any cartographic training (Faby & Koch, 2010; Haklay, et al., 2008; Jobst & Döllner, 2008; Turner, 2006). As a consequence, their resulting products might not transfer their message as effectively as intended.

Although these cartographic products have evolved drastically over time, their main function still remains the same as the one identified in the 1950s: communication of geographic information. However, the possibilities and limitations inherently linked with screen displays can have an impact on how maps can be presented to the user, and thus on how the information is perceived (Cartwright, et al., 1999; Robinson, et al., 1995). Recent technological advancements facilitate and even encourage the use of dynamic and interactive map displays, but is this really beneficial for the user? With the rise of web maps and other digital cartographic displays, a number of authors expressed their concerns.
In 2001, MacEachren and Kraak put forward a number of research challenges in geovisualisation. One of the key challenges they identified was based on the cognitive and usability issues in geovisualisation. They expressed the clear need for methods to assess these cognitive issues, taking into account group differences. In the same year, Slocum, et al. (2001) identified six major research themes related to cognitive and usability issues in geovisualisation, which include the evaluation of the effectiveness of geovisualisation methods and the investigation of individual and group differences.

1.1.2. Cartography and cognition

Cognition includes perception, learning, memory, thinking, reasoning and problem-solving, and communication (Montello, 2002). Montello (2002) presented an overview of the cognitive map-design research during the twentieth century. He noted that the rise of the Geographic Information Systems has decreased the interest in cognitive cartography, although computers could facilitate map design research. Montello also stressed the importance of the work How Maps Work by MacEachren (1995). This book contains an excellent overview of a number of psychological theories and their application on the interpretation process of maps. However, very few authors integrated these valuable theories in user studies on screen maps.

In the beginning of the previous century, a number of psychologists formed the basis for our current understanding of how we perceive and organise a visual scene, which is called the Gestalt Theory (MacEachren, 1995; Matlin, 2002). A part of this theory is based on the idea that humans try to organise or group what they see (such as maps). Furthermore, the scene is perceived as a whole, which is more than the sum of the individual parts. This is consistent with a holistic approach. The Gestalt psychologists (including Wertheimer, 1923) defined a number of ‘laws’ regarding this grouping tendency. For example, the law of proximity states that objects that are close together will form a group; the law of similarity states that similar objects will form a group; the law of experience states that familiar shapes or arrangements will form a group. These laws also contribute to the understanding of, among others, the ‘figure-ground principle’ or the influence of experience on scene perception. They
can thus be used as guidelines on how to design effective maps that facilitate their interpretation (MacEachren, 1995).

When studying a cartographic product (or moving through an environment), the information that is stored in memory (learned) is often called a cognitive map (Downs & Stea, 1997). The cognitive or mental map contains knowledge regarding the spatial layout of certain features, which is stored in memory in the form of a hierarchical structure with nodes and paths (according to Hirtle & Jonides, 1985; Huynh & Doherty, 2007) or by feature and structural information (according to, among others, Kulhavy & Stock, 1996).

Bunch and Lloyd (2006) and Harrower (2007) gave a theoretical overview of the map users’ cognitive processes while working with geographic visualisations, based on psychological theories and previous research. They focus on the limited cognitive load that humans have at their disposal in order to process (interpret) and learn (visual) stimuli. This was also previously examined in detail by Sweller (1988). Harrower (2007) made suggestions for a more effective map design, but these were not evaluated in user tests.

Hegarty, et al. (2010) manipulated the relative visual salience of task-relevant and task-irrelevant information on different maps and investigated the participants’ performance in relation to their domain knowledge. They found that a good design of the display (task-relevant information more salient) facilitated the performance significantly, especially if the participants possessed domain knowledge (Fabrikant, et al., 2010).

Almost ten years after the paper of MacEachren and Kraak (2001), Fabrikant and Lobben (2009) and Montello (2009) pointed out that some research that considers the map users’ cognitive issues has been undertaken (Çöltekin, et al., 2010; Fabrikant, et al., 2008; among others), but that this research issue still needs attention. The visualisation issues of these neo-cartographic products – which have a very high accessibility rate through the Internet – were considered to be a new research challenge on the agenda of the International Cartographic Association (Virrantaus, et al., 2009).
These neo-cartographic products come along with a number of possibilities, such as the visualisation of change over time using different types of animations: motion, change in colour, change in size, (dis)appearing, etc. However, it was found that humans are not as capable in detecting changes in a visual scene (such as maps) as was assumed. In order to be able to detect a change in a certain scene, the attention of the user has to be focussed on the object that changed (among others, Rensink, 2002; Simons & Ambinder, 2005). Cartwright (2012) also listed a number of issues related cartographic products created by amateur map makers, among others regarding their design, maintenance, privacy, and data protection.

As a conclusion, the previous paragraphs learn that since the rise of digital cartographic displays and related geovisualisation methods, a pressing need to understand the users’ cognitive processes has surfaced. Map users have been studied in the past while working with paper maps (see Montello, 2002 and van Elzakker, 2004 for an overview), but in modern cartography this research is needed even more. Map makers need to understand how the end users of the digital cartographic products perceive, read, interpret, and process the visual information presented to them on a screen. Furthermore, due to the accessibility of web maps, more novice map users – who did not receive any cartographic training – will use these maps. As a consequence, it is essential to understand the difference between how experts (in the field of cartography) and novices (people without cartographic knowledge) process cartographic screen maps. In order to improve the effectiveness of screen maps, the user has to be placed in the centre of the design process of the maps.

1.1.3. *Improving the effectiveness of maps: UCD*

*User Centred Design* (UCD) and *Usability Engineering* (UE) are methods that find their origin in the field of software development. By involving the end user in the different phases of the product’s development, issues in the product’s design are identified at an early stage and can thus be solved (with an acceptable cost) (Nielsen, 1993; Rubin & Chisnell, 2008). This iterative process of ‘designing a product’ and ‘evaluating the product’ allows creating more *usable* or *user friendly* products. The ISO standard *Guidance on Usability* (ISO 9241-11) defines usability as “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”
A typical UCD-lifecycle of a product is presented in Figure 1.2.

In the initial phase, before the actual design of the product, the user requirements have to be considered. In order to create an effective product, the characteristics, preferences, and the context of use regarding the end user have to be identified (Nielsen, 1993). After carefully analysing these requirements, a first prototype is built which may not necessarily contain all requested functionalities. Next, this first prototype should be evaluated in a user study. Based on the results of this evaluation, the initial design can be modified. This iterative process has to be repeated until an acceptable design is reached. Even after the release of this ‘final’ design, user feedback is still essential to contribute to the effectiveness of future products (Nielsen, 1993; Rubin & Chisnell, 2008).

Due to the recent evolutions in digital cartography, a connection with software engineering can be identified. van Elzakker and Wealands (2007) stressed the importance of the requirements phase in the development of multimedia cartography. This connection was also acknowledged by Nivala (2007), who studied the familiarity of map makers with the techniques used in usability engineering and their suitability to evaluate (screen) map designs. They concluded that most map making companies are interested in applying UCD, but that they lack the knowledge on how to implement the approach, including the different techniques. Nevertheless, UCD is recently more and more used in the development and design
of geographic applications, such as location based services (among others, Delikostidis, 2011; Haklay & Nivala, 2010; Roth, 2011).

UCD comprises a wide range of techniques, each of which should be used at an appropriate stage of the UCD lifecycle. Some techniques are used for requirements elicitation, before there is any concrete prototype. Other techniques are employed to test an initial prototype (early in the design process). Another group of techniques can be addressed to evaluate the final product before or even after its release. The main aim of these techniques is to identify usability problems in the design of the prototype. However, some techniques also allow studying the users’ problem solving behaviour and cognitive processes in order to identify why something is causing problems for the users. Examples of such techniques are eye tracking and thinking aloud, which could also be used in functional map design. A good overview of these techniques and their appropriate use is described by Nielsen (1993).

Despite the pressing need to evaluate cognitive issues related to digital cartographic products, many thousands of these products have been made accessible on the World Wide Web, without prior evaluation and even without considering the users’ needs. This is partly due to the recent evolutions of Web 2.0, which made it fairly easy for amateur map makers (novices) to create web maps. Furthermore, no user feedback regarding their usability or effectiveness is gathered after their release. As a consequence, still very little is known on how the end users actually read, interpret, and process these screen maps, even when only considering basic designs (without many animations or interaction possibilities).

Still, a large number of screen maps are designed by experts in cartography (or software engineering), whereas the actual end users are novices in these fields. Since novices might interpret these cartographic displays from a different point of view (influenced by background knowledge, experience, etc.), noise and errors in the communication process might occur (cfr. Figure 1.1) (Dodge, et al., 2011). To solve this discrepancy, map designers need to know how these novice map users process the visual stimuli presented on a screen.
1.2. Rationale and synopsis

1.2.1. Research objectives and questions

The considerations and issues mentioned in the previous sections build the main motivation for the work presented in this dissertation. A need for research that evaluates the cognitive issues in geovisualisation, related to different user groups, has been identified for over a decade now, but remains largely unanswered. This need is even reconfirmed by the thousands of maps currently found on the Internet that were never evaluated by any representative user before their release. In order to create effective digital cartographic products in the future, it is essential that experts in cartography – professional map makers – understand how the novice users read, interpret, and store the visual information presented to them. In this context, three main research objectives are identified:

Research Objective 1:
*Improve the effectiveness of (screen) map designs based on the users’ characteristics.*

Research Objective 2:
*Contribute to the understanding of how map users read, interpret, store, and use the presented visual information on screen maps.*

Research Objective 3:
*Investigate the influence of (cartographic) expertise on the map users’ cognitive processes and their limitations while processing the visual information presented on screen maps.*

The first research objective is a general one, which is further specified in the two other research objectives: RO 2 and RO 3. The research mainly focuses on the design of screen maps. However, the outcomes of these studies might also contribute to the effectiveness of map design in general. The second research objective specifies what we want to do in order to improve the effectiveness of the maps. The third research objective specifies the user groups that are considered. These three main research objectives are translated in four, more concrete, research questions:
RQ 1: *How do map users read and interpret the visual information presented on screen maps?*

When a visual stimulus is presented to a user, he will initially read the content. In order to do this, he has to focus his attention on subsequent regions in the stimulus, using a number of guidelines which determine the location of the next region of focus. Next he will figure out what the content means: interpretation of the information (Dodge, *et al.*, 2011; Muehrcke, 1986; Wolfe, 1994; among others). These steps form the essential first phase when working with visual stimuli, such as screen maps. If we can understand how users read and interpret visual information that is presented on the screen, the map makers can key this design to the users’ cognitive processes. This could guide the users’ attention and would facilitate the interpretation process and, by extension, the communication process (Hegarty, *et al.*, 2010; Klippel, *et al.*, 2005; among others).

RQ 2: *How do map users store and retrieve (use) the information that was previously gathered from screen maps?*

The information gathered from screen maps is only useful if it can be applied later on in other situations. This means that the information, which was stored in the memory after the interpretation process, has to be retrieved again. In order to support effective communication, the information has to be presented in such a way that it can easily be stored in memory; otherwise it would be forgotten after its interpretation. Finally, the communication process is only successful if the user can use this information again later on, for example during the interpretation process of other features (Muehrcke, 1986; among others). Therefore, it is essential to understand how this information is retrieved from the users’ memory structures.

RQ 3: *How are the map users’ cognitive processes influenced by deviations in the map image?*

In order to optimise the design of a map, the visualisation of some elements will have to be adapted: symbolisation, use of colours, appearance, etc. As a
consequence, these elements are visualised differently from what the user is familiar with and could distort the user's cognitive processes (reading, interpretation, storing, and retrieval) and thus the communication process (Dodge, et al., 2011; Keates, 1996; Robinson, et al., 1995; among others). However, very little is known regarding the effect that these adaptations/deviations have on the map users. Due to the new tools in the Web 2.0, non experts in cartography are also creating maps on the Internet and these are often not conforming to the established cartographic rules (Cartwright, 2012). Nevertheless, these (even minor) deviations in the map image might influence the cognitive processes of the (novice) map users in a negative way. This research question is closely linked with research objective 3, which considers the cognitive limitations of map users.

RQ 4: How does (cartographic) expertise influences the cognitive processes investigated in the previous research questions?

Map users cannot be considered as one homogeneous group. Individual differences occur, but users can also be categorised in a number of groups, based on certain similar characteristics. The different characteristics between two such user groups could influence how they work with certain products and process information (i.e., Aykin, 1989; Nielsen, 1993; Rubin & Chisnell, 2008). People who design maps typically have a higher level of (cartographic) expertise than the actual end users. However, due to the lack of (representative) user evaluations during the design process, many maps are still designed from the experts’ point of view, which might not be beneficial for novice map users (Dodge, et al., 2011). Expertise might influence, on the one hand, how the maps are read and interpreted by the different map users (see for example Hegarty, et al., 2010). It might, on the other hand, also influence how this information is stored ‘internally’ and consequently how it will be used later on. Finally, when introducing deviations in the map image, expert and novice users might react differently. Nowadays, novice map users also can create cartographic products on the Internet, which are in turn used by expert or novice users. In order to maintain an effective and efficient communication process, experts could introduce limitations or set guidelines for these amateur maps. As a
consequence, it is essential to identify and study the differences between both user groups. This final research question should thus be treated together with the previous ones.

1.2.2. Dissertation outline
The research questions described above are handled in the following chapters. An overview of these chapters – their relations and contents – is presented in Figure 1.3. This figure shows that there is no one-to-one relationship between the research questions and the chapters. Chapters 2-6 correspond to articles which have been published or submitted to international peer-reviewed journals.

In correspondence with the UCD lifecycle, the Chapters 2-4 describe experiments that are conducted using an initial prototype – screen map in this case – with a basic design. In the early stages of user testing, these basic designs are used to rule out as many influencing factors as possible in order to objectively study the issues at hand. In these initial experiments, three different techniques were used (and combined) to find answers on the different research questions. The reaction time measurements were used to test the users’ performance; eye tracking was used to study the users’ cognitive processes; and finally, post-study questionnaires were used to obtain background information regarding the participants. More details regarding these techniques can be found in the corresponding chapters.

Ch 2: Interpreting Maps Through the Eyes of Expert and Novice Users
This chapter – published in the *International Journal of Geographical Information Science* (Ooms, et al., 2012a) – describes an experiment in which the eye movements and reaction times of both expert and novice users are registered. The, mainly statistical, analyses of the obtained data, reveal how both users groups search (and thus interpret) the visual information on the presented screen maps.
Figure 1.3: Dissertation outline
Ch 3: *Analysing the Spatial Dimension of Eye Movement Data Using a Visual Analytic Approach*

The analyses presented in this chapter – published in *Expert Systems With Applications* (Ooms, *et al.*, 2012b) – are an essential addition to the statistical analyses described in Chapter 2. In order to effectively study how participants study a screen map, the data has to be analysed in a spatial reference frame. Consequently, the data have to be analysed with a qualitative – visual – approach that can map the eye movements to the map’s content.

Ch 4: *Investigating the Effectiveness of an Efficient Label Placement Method Using Eye Movement Data*

As a next step in the UCD lifecycle, small variations in the map design are introduced and evaluated. Two label placement options are considered, still in combination with the basic map design. The proposed label placement technique places labels in less optimal positions in comparison to existing techniques. Eye tracking, reaction time measurements, and questionnaires are used to investigate the influence of these less optimal label placements on the (novice) users’ cognitive processes. The results of this experiment are published in *The Cartographic Journal* (Ooms, *et al.*, 2012c).

More complex stimuli are implemented in the experiments of Chapters 5–6. These stimuli contain all types of map elements (points, lines, and areas) with an elaborate symbolisation and correspond to topographic maps which are currently in use in Belgium. Consequently, the tests described in these chapters correspond to a final stage of user testing in the UCD lifecycle. In these experiments, a combination of eye tracking, thinking aloud, and questionnaires was used to investigate the issues at hand. Eye tracking was used to study the users’ cognitive processes during information interpretation; thinking aloud was used to study the users’ cognitive processes during information retrieval (and to obtain insights in the memory structures); and the post-study questionnaires were again used to obtain background information.
Ch 5: *Understanding Expert and Novice Map Users: Reading and Interpretation*

The cognitive processes of expert and novice users related to reading and interpreting the (complex) map content are investigated in detail here. In order to analyse the recorded eye movement data in the spatial context of the map design, a gridded visualisation approach is proposed. This study is submitted to *Cartography and Geographic Information Science* (Ooms, et al., 2012d).

Ch 6: *Listen to the Map User: Cognition, Memory, and Expertise*

After processing and learning the information presented on the screen maps, users can employ this information again later on. Exactly by being able to reuse the previously processed information, they reach effective communication through maps. Therefore, it is essential to understand how users store and retrieve the information and on how much of the information that was presented on the map could actually be used by the map user. The thinking aloud technique was used to study how map users store and retrieve information derived from (digital) cartographic products. The results discussed in this chapter are submitted to *The Cartographic Journal* (Ooms, et al., 2012e).

The last two chapters summarise and discuss the results of the previous chapters. Chapter 7 comprises an elaborate *General Discussion* regarding the initial research questions in combination with the obtained results from the different experiments (and chapters). This chapter also points out some venues for future work. The final chapter (Chapter 8) presents a *General Conclusion* regarding the outcomes of this dissertation.

1.2.3. *Out of scope*

Since we cannot test every aspect of the research objectives and questions presented, a number of critical decisions had to be made throughout the different stages of the research. This means that the work presented in this dissertation is also inherently linked with limitations. In every chapter, these decisions and their justifications are described on a detailed level. Next to these detailed decisions and limitations, the overall research had to be conducted within a fixed framework,
which introduced limitations too. This general framework is presented in the next paragraphs.

First of all, the main objective is to improve the effectiveness of screen map designs. Nowadays, these screens can have a wide variety in dimensions which is closely linked with the devices in which they are integrated. These sizes can range from just a few inches (smart phones), over about 10 inches (tablets and netbooks), or 20-30 inches (notebooks, monitors, and flat screens), to more than 50 inches (touch tables and television screens). In the context of this dissertation, it was decided to focus on only one screen size, which dimensions would correspond to these of a (normal) screen connected to a desktop PC: 21 inches. This screen size was selected because it is most commonly used to visualise web maps.

Besides the screen maps, specific criteria for the participants included in the user tests had to be identified. As mentioned before, two user groups are selected, which are called experts and novices. With experts, we mean the persons who have received cartographic training during their academic education in Geography or Geomatics. As a consequence, these ‘experts’ also possess geographic domain knowledge, such as properties of the subsoil, typical layout of human settlements and infrastructures, geo-physical and geological background knowledge, etc. Furthermore, these experts worked at the Department of Geography at the moment of the test, which means that in their daily job they still worked with cartographic products on a regular basis. The novices, who took part in the study, were students who did not receive any cartographic training and who did not follow any science courses. As a consequence, their interests were not situated in the field of cartography (and science in general). This strict identification of both user groups is necessary to be able to study the influence of expertise across the different user studies.

Furthermore, a selection had to be made regarding the type of maps that were implemented in the study. A very wide range of map types and related cartographic product exists, especially when considering the digital medium. Two main categories of maps are distinguished: thematical maps and reference maps. Reference maps are often used as a base layer on which other (thematic) information is
superimposed. Because of this broad functionality of reference maps, it was decided to implement these in the user studies (with different levels of complexity).

Finally, most screen maps that are found in the Internet today allow the users to interact with them: zooming, panning, turning layers on and off, etc. However, in the experiments that are described in this dissertation the participants could not interact with the stimuli. Only a simulated user interaction was shown to the participants (pan operation). This simulation ensured the controlled nature of the experiment, with a fixed time interval and panning distance. This level of control in the stimuli is necessary to objectively compare (and thus analyse) the obtained datasets.

References


Jobst, M., & Döllner, J. (2008). Neo-cartographic influence on map communication in LBS. In: G. Gartner & K. Rehrl (Eds.), *Lecture notes in geoinformation and
cartography (Vol. Location based services and telecartography II pp. 207-219). Berlin Heidelberg: Springer.


Chapter 2

Interpreting Maps Through the Eyes of Expert and Novice Users

Modified from:

The experiments described in this chapter combine response time measurements and eye movement data to gain insight in the users’ cognitive processes while working with dynamic and interactive maps. Experts and novices participated in a user study with a between user design. Twenty screen maps were presented in a random order to each participant, on which he had to execute a visual search. The combined information of the button actions and eye tracker reveal that both user groups showed a similar pattern in the time intervals needed to locate the subsequent names. From this pattern, information about the users’ cognitive load could be derived: use of working memory, learning effect, etc. Moreover, the response times also showed that the experts were significantly faster in finding the names in the map image. This is further explained by the eye movement metrics: experts had significantly shorter fixations and more fixations per second meaning that they could interpret a larger part of the map in the same amount of time. As a consequence, they could locate objects in the map image more efficiently and thus faster.

Keywords: cartography; eye tracking; usability; cognitive map
2.1. Introduction

Usability Engineering (UE) and User Centred Design (UCD) are well-known themes in the domain of software development. UCD involves the user in the subsequent stages of the product’s development to enhance the usability of the final product. By involving the user in the production process, the effectiveness of the product – or its quality towards the user – improves drastically. ISO 9241-11, Guidance on Usability, defines usability as ‘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’ (Earthy, et al., 2001, p.554). In this context, effectiveness is related to how well a user can accomplish a certain task: the accuracy and completeness. Efficiency is related to how fast a user can accomplish a task: learning time and completion time. Finally, satisfaction is related to the user’s preferences.

A frequently used UE-technique to examine, among others, the layout of user interfaces and websites is eye tracking (e.g., Djamasbi, et al., 2010; Fleetwood & Byrne, 2006; Goldberg, et al., 2002; Jacob & Karn, 2003; Schiessl, et al., 2003). This technique allows ‘tracking’ the movements of the participant’s eyes: his Point of Regard (POR) is registered at a certain sampling rate. From this long list of \((x, y)\) positions, eye movement metrics such as fixations and saccades can be derived. A fixation is a stable POR during a certain time span (at least 80 to 100 ms) and indicates that the user is interpreting the content at that location. A saccade is a rapid eye movement between two fixations, typically completed in tens of milliseconds\(^1\). The velocity of these saccades (up to 500 °/s) is such that the user cannot interpret any content at these moments. A scanpath is a succession of fixations and saccades (Duchowski, 2007; Poole & Ball, 2006; Rayner, 1998).

The use of eye movements in user studies is not a new method. One of the first experiments dates from the end of the 19\(^\text{th}\) century. These initial techniques differ significantly from the ones used today. They were rather invasive with direct contact to the participant’s eyes. The first application of eye movement studies in the field of UE is described in the work of Fits, et al. (1950), who used motion picture cameras to study the movements of pilots’ eyes (Jacob & Karn, 2003). Also in cartography,

\(^1\) For example: A distance of 30 cm on a screen (which is viewed at a distance of 70 cm by the user) is covered in 48.379 ms or 0.0484 s.
the use of an eye tracking method to study the user's attentive behaviour is not new. Inspired by the systematic use of eye movement recordings in research fields that involve graphical communication, such as psychology and art, Jenks (1973) studied the scanpaths of users looking at a dot map. Based on this initial study, a number of follow-up studies were conducted during the 1970s and the first half of the 1980s (e.g., Castner & Eastman, 1984, 1985; Dobson, 1977; Steinke, 1987).

After these first studies, researchers recognised the method's applicability, but concluded that no new knowledge could be derived from it. Consequently, the use of eye movements almost disappeared in cartographic user studies after 1985. Jacob and Karn (2003) give three main reasons why the use of eye movements in usability research did not get widely accepted. First, the technical problems related to capturing the actual eye movements were very challenging in the past and produced rather inaccurate and thus unreliable results. Second, eye trackers produced huge amounts of raw data from which meaningful metrics needed to be extracted. This labour-intensive and often manual data extraction complicated and slowed down the analysis significantly. Third, the interpretation of the extracted data was very difficult. Furthermore, some doubts emerged regarding the link between a person's eye movements and his spatial attention during a visual search: we can all move our attention without moving our eyes (Montello, 2002; Rayner, 1998).

More recent studies, however, show that eye movements are critical to interpret visual information efficiently while performing a complex visual and cognitive task (Duchowski, 2007; Henderson & Hollingworth, 1998). The eye tracking systems have also evolved drastically during the last decades with new and more accurate techniques that have a smaller impact on the participant himself. Moreover, the cost of the eye tracking devices has decreased considerably during this period. The software packages that come with these eye trackers today allow more flexible extractions of meaningful metrics related to fixations and saccades (Duchowski, 2007; Goldberg, et al., 2002; Jacob & Karn, 2003; Poole & Ball, 2006; Rayner, 1998). Not only what a user is looking at but also how long, how often, the length and speed of the saccades, etc. can be discovered. As a consequence, more detailed insight in the user's cognitive processes can be derived from these measurements in comparison with the initial eye movement studies. Even more, during these last
decades psychological research on cognitive processes linked with visual search has received much attention, which resulted in new and more detailed theories regarding cognitive cartography (e.g., Harrower, 2007; MacEachren, 1995; Slocum, et al., 2001)

At the same time, also the maps have experienced a tremendous evolution during this period: from a static, analogue format to highly dynamic and interactive digital maps. In 2001, Kraak defined a rough web map classification based on how the map is or can be used. He distinguished between static and dynamic maps on the one hand and between view-only and interactive maps on the other hand. Kraak (2001, p. 3) noted that ‘The most common map found on the WWW is the static view-only map.’ Today, however, almost every map on the Internet is ‘clickable’ and produces dynamic responses such as animations or videos.

The recent evolutions in cognitive cartography and the improvement of the eye tracking method has resulted in a renewed interest in eye movement research in the field of cartography during the past few years, but so far, only a few studies have been conducted. Brodersen, et al. (2001), for example, investigated the symbology of analogue (paper) topographic maps. Fabrikant, et al. (2008) and Çöltekin, et al. (2010; 2009) recorded participants’ eye movements to evaluate animations and interactive interfaces related to the presentation of maps. These initial eye movement studies (using the improved eye tracking techniques) prove the suitability of eye movement research to investigate how users perceive these highly dynamic and interactive maps of the current digital era. Montello (2009) also recommended the recording of eye movements as a future method for cognitive GIScience research.

The goal of this chapter is to extend these eye movement studies to obtain insight in the user’s cognitive processes while working with these dynamic and interactive maps: construction of the mental map, cognitive load, learning effect after interactions, etc. These observations can be compared with the Cognitive Load Theory, that Harrower (2007) and Bunch and Lloyd (2006) described in relation to the interpretation of map animations.
The Cognitive Load Theory distinguishes between two main memory types: working memory and long term memory. The working memory is used when new information is processed. It cannot contain large amounts of data or store it for a long time. In this case, the user’s long term memory is addressed. Consequently, information has to be transferred from the working memory to the long term memory. This learning task is achieved by linking the current (active) information in the working memory with knowledge and skills already stored in the long term memory. When necessary, information stored in the long term memory can be retrieved and used again in the working memory. The Cognitive Load Theory describes three types of cognitive load that have an influence on the working memory and the learning task. The intrinsic cognitive load is related to the complexity of the visual information: complex information is more difficult to process and results in a higher cognitive load. The second type of cognitive load, the extraneous cognitive load, will increase due to distractions or a poor representation of the information. Finally, the germane cognitive load is closely linked to the learning task itself: a part of the working memory is used to link the active information (in the working memory) to previous knowledge in the long term memory so that the active information can be passed on to the long term memory (Bunch & Lloyd, 2006; Harrower, 2007).

What is more, different users may interpret and process the same information in a different way. Consequently, map users cannot be considered as one homogeneous group but as different categories of users and individual user differences have to be taken into account. Different user characteristics of interest are: gender, age, experiences, background knowledge, etc. (Aykin, 1989). Important and interesting differences between users are the background knowledge and the level of experience they have with the topic under investigation, maps in this case (MacEachren, 1995; Nielsen, 1989). Since the interpretation process is closely linked to the structure and use of the user’s memory, it is influenced by previous knowledge. As a consequence, designers of user studies often differentiate between expert users (high level of experience) and novices (low level of experience) (Duchowski, 2007; Nielsen, 1993; Rubin & Chisnell, 2008).

This chapter gives an overview of the results from an experiment conducted both by novices and experts. Since the same stimuli – screen maps in this case – are
presented to two different user groups, the study has a between user design (Duchowski, 2007; Nielsen, 1993). The comparison of the results allows detecting whether experts (persons with cartographic training and experience in interpreting maps) can interpret maps more efficiently. A combination of response time measurements and eye movement data is used to obtain insight in the participants’ cognitive processes while working with dynamic and interactive maps.

2.2. Study design

2.2.1. Participants

The participants of the expert group were at the moment of the study employed at the Department of Geography at Ghent University. They obtained at least a Master degree in Geography or Geomatics and received, both theoretical and practical, cartographic training. In their daily job they use paper and digital maps on a regular basis. Consequently, the expertise of this group is two-fold. On the one hand, they have a substantial level of background knowledge of cartographic syntax and semiotics, on the other hand, they are highly experienced in working with paper and digital maps.

The group of novices were Bachelor students of the Faculty of Psychology and Educational Sciences at Ghent University. None of them received any previous training in cartography. The group of expert users counted 16 participants, whereas the group of novices 15. The data of one novice participant was omitted because this participant was familiar with a number of regions in the stimuli. All participants cooperated on a voluntary basis and were Dutch-speaking (the language in which the test was presented).

2.2.2. Stimuli

Twenty demo-maps were presented to each participant in a random order on a screen. An example of a demo-map is presented in Figure 2.1a. The design of these maps was very basic and controlled. The next sections explain the reasons for this specific design.
First of all, the design had to be homogeneous within one map, without any deviating regions in it. Secondly, the design of all maps had to be very similar in order to avoid users being distracted in any way. The cognitive load related to processing the information on these maps needed to be limited, equal for all maps, and equal for all participants. In this way, the design prevented that certain users considered a specific area on a map more interesting or that a certain map would be more interesting than the others.

In order to keep the extraneous cognitive load limited and equal for all maps, the same simple background image was used for every map. The simplest background image of a map consists of a small number of polygons (for example three), filled with a non-obtrusive pastel colour. These polygons can present, for example, countries or thematic regions. Adding extra objects to this background (such as lines or more polygons) would increase the complexity of the map, resulting in a higher intrinsic cognitive load for the user. Using one or more striking colours could distract the users, resulting in a higher extraneous cognitive load. Too many or different map elements across the twenty trials would distract the participant from the actual task. The complexity of the information presented on the foreground also needed to be limited and balanced across the twenty maps. The number of elements, types of elements, and their distribution on the map contribute to the intrinsic cognitive load of the user: the complexity of the visual information. Therefore, only point objects and their associated labels were depicted, which may represent cities or points of interest.
The number of labels depicted was similar for all twenty maps. The actual visualisation of these objects has an influence on the extraneous cognitive load. In order to keep this load low, all labels were visualised in the same colour (black) and only a limited level of hierarchy in the labels was present. Some labels were presented in capitals, but most were depicted in a normal font. The names in the task were rather short (consisting of, on average, five characters) and not written in capitals. What is more, the distribution of cities (points) and their names originated from existing regions in order to present realistic situations to the participants. However, these regions were selected in such a way that none of the Belgian participants was likely to recognise one of the regions. Familiarity of a participant with a certain region would mean that the participant has already stored some information about that region in his long term memory. The germane cognitive load for this participant is influenced differently when viewing the familiar regions: he can make a direct link between the working memory and the long term memory.

After a fixed interval of 50 s, a user interaction was simulated: a horizontal pan operation to the right. As a consequence, a part of the initial view remained visible on the left side of the screen and a new part of the map emerged on the right. This fixed time interval ensures that all users will look at the initial map for the same amount of time. Figure 2.1b depicts such a view after the interaction.

It was chosen to use a basic map design in the test to avoid as many influencing, possibly confounding factors as possible. This map design also implies that the value of expertise, related to the expert users, is limited. Using very complex maps with multiple layers of information might be prejudicial for the novice map users because they do not have the same background knowledge and level of experience. In extension to this study, new experiments can be designed with an increased level of map complexity, related to the objects themselves or to the visualisation of the objects. This way it can be discovered which elements or what level of complexity has a profound impact on the different types of users: novices and experts. However, this extension is beyond the scope of this chapter.
2.2.3. Procedure
On the right side of each map, a list of five names was presented (see Figure 2.1a). The participant was asked to locate these names in the map image and to push a button each time a name was located. The participant could use any of the buttons at the bottom of the joystick in order not to distract him from the task or to discriminate between left and right handed people. Each of these button actions resulted in a response time measurement. In order not to disturb the cognitive processes of the participants, they were free to choose in which order they wanted to locate the five names. According to a previous (unpublished) study by the authors it was found that most participants could locate these five names within 50 seconds. Therefore, the horizontal pan operation was simulated on the screen after this time interval. As a result of this design users can anticipate when and how the interaction will occur, which is considered a positive element: when users perform an interaction ‘in real life’ they also know in advance when and how this will occur. Unexpected changes in the map view, both in terms of time and direction may confuse the user.

During the simulated pan operation, the list of five names also changed (see Figure 2.1b). Three of the initial names reappeared in the new list (on a different position) and two were new. The user had to locate these names in the map image again and push one of the two buttons at the bottom of the joystick to confirm he found one. When all names were located, the user could end the trial by pushing one of the buttons on top of the joystick.

Before the actual start of the test, two demo trials were presented to the participant. The goal of these demo trials was two-fold. First, the participant could practise on the assignment allowing him to ask any additional questions before the start of the actual experiment. Second, he was able to familiarise with the background and general structure of the maps used during the actual tests.

The task described above corresponds to a visual search on an image. A visual search is a very realistic and natural operation for a map user. He is trying to find locations, such as cities, on a map of which he only knows the name. When exploring the surroundings of these locations, the user will use a pan operation to visualise a
larger part of the map. After the pan operation, the user tries to orient and explore
the map by locating new names and relocating known names from the view before
the interaction.

2.2.4. **Apparatus**
The study was conducted in a controlled environment: the Eye Tracking Laboratory
in the Faculty of Psychology and Educational Sciences (Ghent University). An
EyeLink 1000 eye tracking device (SR Research, Ontario, Canada) was used to
record the participant’s eye movements during the study. This system can sample the
participant’s Point of Regard once every millisecond. Each participant received a
joystick (Microsoft Sidewinder Plug-and-Play Game Pad) from which two button-
clusters could be used. The stimuli were presented on a 21 inch monitor with a
resolution of 1280 by 1024 pixels.

2.2.5. **Recordings**
During this study, different types of recordings were obtained and combined. First,
response time measurements were obtained from the button actions. With these
button actions, the participant indicated that he found a label. However, a user
could push the button without finding a correct name. Therefore, these absolute
response times were compared with the eye movements. If a user was not focussing
on a name from the list, the measurement was removed from the dataset. In the
analyses, the relative response times are used: the time interval between locating
two subsequent names. Hence, locating a wrong label influences the relative
response time of the next label. Consequently, these measurements, related to the
subsequent label, were also removed from the dataset.

2.3. **Results**
2.3.1. **Response time measurements**
The relative response time measurements indicate how fast a certain user could find
each subsequent label. The overall (all labels, all users) mean \( M \) relative response
time was 5.472 seconds \( (SD = 1.605 \text{ s}) \). A one way ANOVA shows that the experts
were significantly faster at finding the subsequent labels \( (M = 5.227 \text{ s}, SD = 1.433 \text{ s}) \)
than the novices \( (M = 5.595 \text{ s}, SD = 1.673 \text{ s}) \), with \( F = 7.056 \) and \( P = .008 < .01 \).
When considering only the time intervals before the simulated interaction, a marginal significant difference could be found in the response time measurements ($F = 3.000; P = .084$). During this time interval the experts’ mean relative response times were shorter ($M = 5.619$ s, $SD = 1.462$ s) than those of the novices ($M = 5.965$ s, $SD = 1.648$ s). After the simulated interaction, the experts were also significantly faster at locating the names than the novices ($M_{exp} = 4.836$ s, $SD_{exp} = 1.297$ s, $M_{nov} = 5.235$ s, $SD_{nov} = 1.622$ s, $F = 4.594, P = .033 < .05$).

In Figure 2.2, the mean response times (with their associated 95% confidence intervals) are represented for each subsequent label, both before and after the simulated interaction. These graphs indicate that the mean response time measurements for the experts were always lower than those of the novices, but follow the same pattern. This pattern can also be derived from the actual values in Table 2.1. Both before and after the interaction, the shortest time interval was linked to the second label. After this label was found, the relative response times grew with each subsequent label, also both before and after the interaction. The time needed to locate the first label before and after the interaction was always higher than for the second label. Furthermore, the response times before the interaction were always longer than after, compared with the corresponding label (e.g., Label 3 before vs. Label 3 after).

![Figure 2.2: Mean values (with 95% confidence interval) of the relative response times for locating the subsequent labels](image_url)
Table 2.1: Mean values for the time intervals (in s) for locating subsequent labels and their statistical comparison

<table>
<thead>
<tr>
<th></th>
<th>Before</th>
<th></th>
<th>After</th>
<th></th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>F</td>
</tr>
<tr>
<td>Label 1</td>
<td>5.934</td>
<td>1.311</td>
<td>4.576</td>
<td>1.203</td>
<td>11.657</td>
</tr>
<tr>
<td>Label 2</td>
<td>4.969</td>
<td>1.315</td>
<td>4.316</td>
<td>.957</td>
<td>3.218</td>
</tr>
<tr>
<td>Label 3</td>
<td>5.537</td>
<td>1.283</td>
<td>4.564</td>
<td>1.179</td>
<td>6.234</td>
</tr>
<tr>
<td>Label 4</td>
<td>5.815</td>
<td>1.264</td>
<td>5.053</td>
<td>1.145</td>
<td>3.987</td>
</tr>
<tr>
<td>Label 5</td>
<td>5.840</td>
<td>1.945</td>
<td>5.670</td>
<td>1.575</td>
<td>.093</td>
</tr>
</tbody>
</table>

*a. experts*

<table>
<thead>
<tr>
<th></th>
<th>Before</th>
<th></th>
<th>After</th>
<th></th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>F</td>
</tr>
<tr>
<td>Label 1</td>
<td>5.954</td>
<td>1.805</td>
<td>4.756</td>
<td>1.265</td>
<td>11.722</td>
</tr>
<tr>
<td>Label 2</td>
<td>5.728</td>
<td>1.991</td>
<td>4.619</td>
<td>1.504</td>
<td>7.894</td>
</tr>
<tr>
<td>Label 3</td>
<td>5.834</td>
<td>1.392</td>
<td>4.860</td>
<td>1.082</td>
<td>12.187</td>
</tr>
<tr>
<td>Label 4</td>
<td>6.069</td>
<td>1.685</td>
<td>5.789</td>
<td>1.469</td>
<td>.631</td>
</tr>
<tr>
<td>Label 5</td>
<td>6.195</td>
<td>1.313</td>
<td>6.152</td>
<td>2.074</td>
<td>.012</td>
</tr>
</tbody>
</table>

*b. novices*

A one way ANOVA shows for both user groups a highly significant difference in the mean response times before versus after the interaction ($F_{exp} = 16.071$, $P_{exp} < .001$; $F_{nov} = 19.372$, $P_{nov} < .001$). The last columns of Table 2.1a and 2.1b show the results from the statistical comparison between the corresponding intervals. These $P$-values indicate that, for both user groups, only the three first corresponding intervals are significantly different. For the experts, the intervals related to the fourth label showed a marginal significant difference, whereas no significant difference was detected for the novices. The last corresponding time interval (Label 5) showed no significant difference for the two user groups. According to Figure 2.2, both the expert and the novice map users show a steep increase in the response times for each subsequent label after the interaction, which is not the case before the interaction. Furthermore, this increase is more pronounced for the novices than for the experts. As a consequence, the fourth and fifth corresponding intervals show no significant difference for the novices, but a marginal significant difference is detected in the fourth corresponding interval of the experts.
2.3.2. Fixation duration

Longer fixation durations can give insight into two elements: the user has more difficulty with interpreting the visual input or the user considers the visual input more interesting. In this study, the latter explanation is not considered. Because of the basic design of the maps, it can safely be assumed that both user groups look at all maps with the same level of interest. The longer fixations measured during this study thus indicate that the user needs more time to process the (visual) information. The overall (all labels, all users) mean fixation duration was .255 s (SD = .086 s), with .249 s (SD = .080 s) for the expert group and .265 s (SD = .094 s) for the novices. A one way ANOVA shows that these eye movement metrics differ significantly ($F = 44.176, P < .001$). Before the interaction, the mean fixation duration linked to the expert group was significantly shorter ($M = .249$ s, $SD = .081$ s) than these linked with the novices ($M = .260$ s, $SD = .090$ s), with $F = 9.315$ and $P = .002$. After the interaction, the experts also had significantly shorter fixations ($M = .248$ s, $SD = .077$ s) than the novices ($M = .270$ s, $SD = .097$ s), with $F = 39.858$ and $P < .001$.

These measurements also indicate that, for the novices, the mean duration was numerically longer after the interaction than before. A one way ANOVA shows that this difference is significant ($F = 6.119, P = .013$). For the expert users, however, the mean fixation durations were very similar during both intervals ($F = .123, P = .726$). An overview of the fixation durations related to finding the ten labels is listed in Table 2.2. The distribution of these measurements is depicted in Figure 2.3 and reveals that the mean fixation durations diverge between the two user groups with each subsequent label that was found, both before and after the interaction. The interval associated with locating Label 1 after the interaction shows an anomaly in this trend: the fixation duration of the novices was very high, but no deviation was noticed in the expert group.
Table 2.2: Mean values for the fixation durations (in s) for locating subsequent labels, both before and after the interaction

<table>
<thead>
<tr>
<th>Label</th>
<th>Before</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>After</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Experts</td>
<td>.260</td>
<td>.100</td>
<td>.240</td>
<td>.092</td>
<td>.242</td>
<td>.067</td>
<td>.253</td>
<td>.080</td>
<td>.244</td>
<td>.071</td>
</tr>
<tr>
<td>Novices</td>
<td>.257</td>
<td>.100</td>
<td>.263</td>
<td>.095</td>
<td>.252</td>
<td>.100</td>
<td>.272</td>
<td>.081</td>
<td>.265</td>
<td>.101</td>
</tr>
</tbody>
</table>

Figure 2.3: Mean values (with 95% confidence interval) of the fixation durations related to each subsequent time interval

2.3.3. **Fixation count**

The number of fixations a user can have per second is closely related to the duration of the fixations. If the duration of the fixation is very long, the number of fixations per second will decrease. However, the combined results of the fixation count and duration provide a better insight in the user’s cognitive processes. Other elements that have an influence on the fixation count are the saccades. Shorter saccades may increase the number of fixations per second.

The mean fixation counts for each subsequent label are listed in Table 2.3 and their distribution is depicted in Figure 2.4. The overall (all labels, all users) mean fixation count was 3.658 fixations per second ($SD = .762$ fix/s). Experts had a higher count ($M = 3.694$ fix/s, $SD = .750$ fix/s) than the novices ($M = 3.605$ fix/s, $SD = .775$ fix/s). A one way ANOVA shows that the measurements of both user groups differ significantly ($F = 16.716, P < .001$). As can be derived from Figure 2.4, the experts...
had significantly more fixations per second (M = 3.713 fix/s, SD = .736 fix/s) than the novices (M = 3.597 fix/s, SD = .754 fix/s) before the simulated interaction, with F = 14.726 and P < .001. The same goes for after the interaction: a significantly higher fixation count for the experts (M = 3.675 fix/s, SD = .764 fix/s) than for the novices (M = 3.613 fix/s, SD = .796 fix/s), with F = 3.940 and P = .047 < .05.

Table 2.3: Mean values for the fixation counts (fix/s) for locating subsequent labels, both before and after the interaction

<table>
<thead>
<tr>
<th>Label</th>
<th>Before</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.654</td>
<td>.834</td>
<td>.702</td>
<td>.736</td>
<td>.740</td>
<td>.795</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.617</td>
<td>.797</td>
<td>.722</td>
<td>.730</td>
<td>.849</td>
<td>.605</td>
</tr>
</tbody>
</table>

Figure 2.4: Mean values (with 95% confidence interval) of the fixation counts (fix/s) related to each subsequent time interval

Similar as with the fixation durations, the number of fixations per second diverged between novice and expert users with each subsequent interval. Equally, the interval associated with locating Label 1 after the interaction of the novice group showed a deviation in the measurement: a much higher number of fixations per second. When comparing the measurements within both user groups, no significant difference could be detected between the intervals before and after the interaction (F_{exp} = 1.902, P_{exp} = .168 > .05 and F_{nov} = .223, P_{nov} = .637 > .05).
2.3.4. Fixation distribution

Analysing the eye movement data qualitatively and visually reveals spatial patterns in the search behaviour of users. The location of fixations indicates which part of visual stimuli (map in this case) is of most interest to a user at a certain moment. Since this experiment focuses on two types of users, differences in these patterns might be noticed. In Figure 2.5, a number of intervals are depicted for each user group, presented by a grid of nine areas of interest (AOIs) that cover the whole map image. The number of fixations in each AOI during the mentioned time interval is presented in percentages. The darker the cell, the more fixations were registered. With this figure, the locations of fixations before the interaction can be compared with the corresponding intervals after the interaction, on the one hand, and between the two user groups, on the other hand. Furthermore, evolutions in the user’s search behaviour might be visible in the subsequent time intervals.

Before the interaction, the distribution of the fixations was very similar between the novice and expert users. During the first second of the map’s display more than 50% of the fixations were located in the middle of the map due to the drift correction: the user had to look at a fixed point in the middle of the screen and the deviation was measured. This was executed after each trial to control the quality of the current eye tracker’s calibration. During the next few seconds, the grid of AOIs shows a higher concentration of fixations in the top row. This indicates that the users tended to start their search in the upper part of the map. Only after three seconds, the fixations seem to be more homogeneously spread over the different AOIs and thus over the map image. During the remaining intervals, the AOIs in the middle seem to have the lowest count whereas the AOIs on the top right corner always have a rather high count. A possible explanation is the list with names located on the right side of the map, but outside the map image. When starting to search for the names, some of the fixations in the scanpath are located in the closest AOI.

After the interaction, no drift correction was executed, resulting in a more homogeneous spread of the fixation during the first second. Similarly as before the interaction, both user groups started scanning the upper part of the map whereas they focussed only in a later moment on its lower part as the higher concentration of fixations in the top row of AOIs reveals. However, in contrast with before the
interaction, the novice users systematically had more fixations in the left column of the AOIs, thus on the left side of the map. The fixations of the expert users were concentrated in the middle and right column of AOIs and much less in the left one.

<table>
<thead>
<tr>
<th></th>
<th>Novices</th>
<th>Experts</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1s</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1-2s</td>
<td>27</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>2-3s</td>
<td>18</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>3-4s</td>
<td>19</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>4-5s</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>5-10s</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>10-20s</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>0s-end</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>32</td>
</tr>
</tbody>
</table>

Figure 2.5: Distribution of the fixations of both user groups on the map image (in percentage) during a fixed time interval

2.4. Discussion

The combined results of the response time measurements and eye movement recordings allow detecting how efficient and effective a user can perform a certain task related to map use. The response time measurements give insight in how fast the participant finds a name and the position of the eye movement at that moment.
reveals if he found a correct one. Furthermore, eye movement metrics (fixation duration and fixation count) give more detailed insight in the users’ cognitive processes while performing a visual search. This insight explains the results of the response time measurements related to the Cognitive Load Theory. The ‘between user’ study design allows differentiating between both user groups: does the background knowledge and experience have an influence on the user’s cognitive processes?

For both user groups, the same trend was visible in the response time measurements for each subsequent label. The smallest time interval was associated with the second label (both before and after the interaction), not with the first label. This first longer interval may be explained by the orientation process. The moment when a map is presented to the participant, he has to orient the map and the list with names before he can start searching on the map. This orientation process will be longer in the initial view because the map is completely new to the user. After the interaction, however, only part of the map is new and thus the orientation time is reduced. After the second label, the response times increased with each subsequent label, which may indicate an equal increase in the cognitive load. On the one hand, the user is trying to locate a current label. This task requires the working memory to process the active visual stimuli and link this information to the previously gathered information (stored in the long term memory). On the other hand, the position of previously located labels is being processed. This process also addresses the working memory which may increase the germane cognitive load in order to be able to transfer this information to the long term memory. With each new label, longer response times were measured, which may suggest an increase in the (germane) cognitive load. The reduced response times after the interaction may be explained by a reduced germane cognitive load. Users may find it easier to create links with the long term memory. However, the response times related to the two last labels after the interaction showed a steep increase, which was even more pronounced for the novice users. This steep increase suggests an equally steep rise of the novice users’ cognitive load.

The shorter response times after the interaction may indicate a learning effect: less time was needed after the interaction because the cognitive load of the user is lower.
A part of the map the user was already familiar with, remained visible after the interaction: a mental map for this part of the map was constructed and stored in the long term memory. After the interaction, this mental map had to be completed with only a small part of the current view that could be linked with the former view. These elements suggest a reduction of the germane cognitive load after the interaction. This implies that the user could invest more of his working memory in processing the current visual stimuli. This investment could explain the significantly faster retrieval of the labels in the second view.

However, the reaction times of the experts were significantly shorter than these of the novices. Although these differences were less than a second, they have to be interpreted in the context of the study design: locating a name on a basic map. These differences may be explained by the eye movement metrics, providing insight in the users’ cognitive processes while working with these dynamic and interactive maps. First, the duration of a fixation may indicate the degree of difficulty experienced when interpreting a certain (visual) content. The results from the experiment show that experts had significantly shorter fixations than the novices, both before and after the interaction. This suggests that experts had it easier to interpret the map’s content than novice users. The background knowledge and experience of the experts might be the cause of this increased efficiency: previous knowledge and habit facilitate interpreting complex visual stimuli. Second, the number of fixations per second is considered. During a fixation the user interprets the visual input. The more fixations the more visual input the user can interpret. Consequently, the user can interpret more content, or a larger part of the map. The number of fixations per second is closely related to the duration of the fixation. The results of the experiment show that experts had significantly more fixations per second than novices, also both before and after the interaction. As a consequence, experts could ‘scan’ or interpret a larger part of the map image in the same amount of time and they can locate the names on the map more efficiently, resulting in shorter response time measurements.

The significantly longer fixations of the novice group after the interaction were mainly caused by the deviation of the measurement during the first time interval after the interaction. A peak is noticed in the measurements, which is not the case
for the expert group. The results from the fixation counts show the same peak during this interval for the novices and not for the experts. Normally, longer fixations are typically associated with fewer fixations per second, but this is not correct for this interval. Since this interval is situated right after the simulation, it could be assumed that the novice map users might have been distracted of the presented pan operation. However, this simulation does not seem to have an influence on the response times of the experts.

The qualitative and visual presentation gives more insight in the distribution of the users’ fixations, their evolution over time and differences between both user groups. The grid of nine AOIs with the associated fixation count (in percentages) gives an overview of the regions on the map that are of interest at a certain moment in time. Both before and after the interaction, both user groups start interpreting the upper part of the map. Gradually more fixations are found in the middle and the lower part of the map. Before the interaction, there was not much difference in the distribution of the fixations between the novices and experts. After the interaction, however, the novices tended to have more fixations on the left side of the map than the experts. The left side corresponds to the part of the map that was already visible during the initial view (before the interaction). The novice users seemed to be more attracted to this familiar part of the map whereas the experts had a remarkably low percentage of fixations in this region. This could be explained by a better structured cognitive map of the expert group, which would make it easier to determine whether a certain label is within the overlapping part after the interaction or not. They only searched on (‘fixated’) the left part of the map if a label (from the list) is located in this region. Novices experienced more difficulties in determining whether a label was already visible on the initial view or not, and consequently searched more often in the left part of the map.

### 2.5. Conclusion and future work

The analyses described in this chapter reveal how users look at, interpret and search on maps, which is essential information to understand the users’ cognitive processes while working with these maps. This understanding is crucial to be able to create more ‘user friendly’ or effective maps in the future, especially since animations and
dynamic interactions are increasingly being added to the interface of maps on the Internet. Recent technologies allow and even support this evolution, but the limits of the end users’ cognitive processes need to be considered. Users who are not that familiar with working on maps might not benefit from all the interaction and animation possibilities since they need more time to interpret the content. These technical possibilities could also be used to differentiate the map content according to the type of user. However, up till now very little practical knowledge is gathered about these cognitive processes related to (dynamic and interactive) screen maps.

The experiment described in this chapter is part of a larger study to obtain more detailed insight in the cognitive processes of map users while working with dynamic and interactive maps. Since maps are essentially spatial objects, the statistical analysis will be extended with a more detailed visual analysis of the eye movement. The visualisation of the users’ scanpaths might reveal patterns in the orientation and/or search behaviour, as well as influencing factors. Furthermore, new tests are planned, using topographic maps that deviate even more between different user groups.

References


Chapter 3

Analysing the Spatial Dimension of Eye Movement Data Using a Visual Analytic Approach

Modified from:

Conventional analyses on eye movement data only take into account eye movement metrics, such as the number or the duration of fixations and length of the scanpaths, on which statistical analyses are performed for detecting significant differences. However, the spatial dimension in the eye movements is neglected, which is an essential element when investigating the design of maps. The study described in this chapter uses a visual analytics software package, the Visual Analytics Toolkit, to analyse the eye movement data. Selection, simplification, and aggregation functions are applied to filter out meaningful subsets of the data to be able to recognise structures in the movement data. Visualising and analysing these patterns provides essential insights in the user’s search strategies while working on an interactive map.

Keywords: map design; visual analytics; eye movements; mental map
Chapter 3

3.1. Introduction
In the field of Human Computer Interaction (HCI), usability studies are conducted throughout a product's design lifecycle to improve its design and functionalities towards the end user. Many types, structures, and techniques exist among these usability studies. The selection of the 'best' method to tackle a certain problem or question depends on a number of factors: stage in the product's development, qualitative or quantitative output, type of (research) question, among others (Nielsen, 1993; Rubin & Chisnell, 2008). ISO 9245-11, a well known standard in the field of User Centred Design (UCD), describes usability in terms quality of use or the efficiency, effectiveness, and satisfaction with which a user achieves a certain goal on a specific system, such as a map (Bevan, 1995; ISO, 1994).

The goal of a good map design is to present the information towards the user in an effective way: the user has to be able to interpret the (spatial) information correctly, but also efficiently. This latter element is related to how the user interprets, processes, and stores 'internally' the information presented to him. In order to be able to improve the design of a map towards the user, it is thus essential to obtain insights in this cognitive or mental map (Downs & Stea, 1977; Montello, 2002).

From the long list of possible usability evaluation methods, the eye tracking technique is considered to be the most suitable to obtain these insights in an objective way. Furthermore, throughout the entire study the user's Point of Regard (POR) is registered, which is the location where he was looking on the screen at a certain timestamp (Duchowski, 2007; Rayner, 1998). Most usability studies focus on quantitative parameters: number of fixations, duration of the fixations, length of saccades, among others (Jacob & Karn, 2003; Poole & Ball, 2006). However, these parameters do not take into account the spatial context of the data, which is essential when dealing with maps and their design. Moreover, most software packages accompanying the eye tracking devices are not fully suitable to study this spatial dimension of the eye movement data. Consequently, a more qualitative and visual analytic approach is needed to be able to detect patterns in the user's behaviour, and thus in his scanpaths while working with these maps. In a number of recent studies the eye tracking method has been used to study the design of maps.
and their usability: their symbology (Brodersen, et al., 2001), map animations (Fabrikant, et al., 2008) and the design of the map interface (Çöltekin, et al., 2009).

But eye movement data show similarities with other kinds of movement data: long list of locations \((x,y)\) at a certain timestamp \(t\). A number of software packages exist which are able to visualise and analyse this type of movement data and which may thus also be able to handle the eye movement data or the scanpaths as well. The Visual Analytics Toolkit (also known as CommonGIS) is such a software package. Its suitability to visualise the eye movement data in a meaningful way is already briefly demonstrated in the work of Fabrikant, et al. (2008). The Visual Analytics Toolkit is developed at the Fraunhofer Institute IAIS (Sankt Augustin, Germany). Its functionalities – including the aggregation and clustering of the movement data – are described in a number of articles: e.g. Andrienko, et al. (2007), Andrienko and Andrienko (2011).

The goal of this chapter is thus to get insights in the users’ cognitive processes while working with dynamic and interactive maps using the eye tracking technique. Spatial patterns in the user’s search strategies on these maps hold information on how users process, store and analyse the information presented to them. The Visual Analytics Toolkit is used to select, summarise and aggregate the massive amount of eye movement data obtained during the user study. The design of this study is described in Section 3.2. The visual analyses are described and discussed in Section 3.3. A conclusion related to these results is presented in Section 3.4.

### 3.2. Study design

The tests were conducted in the Eye Tracking laboratory of the Department of Experimental Psychology at Ghent University, Belgium. This laboratory is equipped with an Eye Link 1000 device from SR Research (Mississauga, Ontario, Canada) and samples a person’s POR at a rate of 1,000 Hz (or once every ms). The movements from one eye only are recorded during the tests. The recorded eye movements of 14 subjects were analysed. All participants were students and most of them studied courses at the Faculty of Psychology and Educational Sciences, Ghent University.
During the tests, the participants were asked to locate five names on a map, which were listed on the right side of the actual map. As a consequence, the participant had to perform a visual search on the map. This is an operation which is done frequently when a user employs a map: he is trying to find the location of a certain object on the map (such as a city, landmark, or crossroad). The user has to orientate and interpret the map content and subsequently search for the location of the target object. The task the users had to execute during this study thus allows studying patterns in the search behaviour which might give insights in how users orientate, process and store the (spatial) information on the map.

Each demo-map had the same simple background with point objects (symbolising cities) and associated name labels. After 50 s the map image was translated horizontally over a fixed difference, simulating a pan operation. This simulation has a duration of one second. The list with the five names had also changed during this translation: two new names are displayed and three which were already in the former list (but on a different location in the list). Again the user had to locate these five names in the map. It was decided to display two different lists of five names to make sure that the user could not track the relative position of the labels in the list during the transition. Since the list is visible during the entire trial, the user did not had to remember all five names reducing the mental workload. The duration of the display of the initial view (50 s) is based on reaction time measurement conducted earlier. In total, 20 demo-maps were displayed to the participants in a random order. An example of such a demo-map is depicted in Figure 3.1.

During the tests, the eye movements of the participants were recorded. Furthermore, the participants were asked to indicate when they found a label by pushing a button. To avoid distracting the users too much from their main task (the visual search) they had to push the same button for each label. Otherwise they might look down to the joystick to look for the button they had to push for a specific label, which would disturb both the time measurements and the recordings from the eye tracker. The combination of the time measurements form the button actions and the location where the user was looking (derived from the eye movement recordings) allows identifying if and which label was found.
The obtained eye movement data is imported into The Visual Analytics Toolkit, where the eye movements (scanpaths) are treated as if it were trajectories from, for example, GPS tracks. In the Visual Analytics Toolkit each point corresponds to a stop and thus to a fixation. The saccades are visualised as lines between the corresponding fixations. In total 77,069 points were loaded into the Toolkit, based on which the trajectories or scanpaths are visualised. On average, 344 points or fixations are registered per trial, with a minimum of 123 and a maximum of 719. The average number of fixations per user over all trials is 5,505.
3.3. Results and discussion

Since eye movement studies result in a vast amount of data, the obtained view is totally overcrowded when visualising the movement data of all participants, on all demo-maps and during the whole time interval, even when the scanpaths are visualised with a transparency of 50%. This is depicted in Figure 3.2. Consequently, it is not possible to detect any patterns and draw any conclusions from this dataset as a whole. Filtering techniques based on certain parameters are needed to select a meaning subset of the scanpaths. The possibilities of the filtering techniques in The Visual Analytics Toolkit and their results are described in the Section 3.3.1.

Figure 3.2: Overcrowded visualisation if all eye movements are selected

3.3.1. Filtering of the movement data

The parameters on which the data can be filtered are linked with the scanpaths’ attributes, on the one hand, or with a time interval, on the other hand. In Figure 3.3, filtering techniques on both types of parameters are depicted and the scanpaths are visualised with a transparency of 50%. Figure 3.3a presents all scanpaths of one participant only, on all demo-maps. Figure 3.3b depicts the scanpaths of all participants for all demo-maps but only during the first ten seconds of each trial. It is clear that more restricted selection techniques are needed to be able to detect any patterns in the movement data: shorter time intervals and/or a combination of several filtering parameters.
In Figure 3.4, a time series of the movement data is presented with subsequent time intervals of 0.50 s. This time series allows visualising and analysing the evolution of the scanpaths over time: how do users scan the map and is there a difference in the search pattern before and after the simulated pan operation?

From this time series it can be derived that nearly all users start looking at the map near its centre. Next, almost all scanpaths are directed towards the list with the five names, which is interpreted during a certain amount of time. After the first second, most users have started searching for the names on the map. By visualising the actual scanpaths it becomes clear that most users start searching for the names in the upper part of the map. After two seconds, the search patterns are more homogeneously spread over the map which still continues before the simulated interaction. This simulation has duration of one second and occurs during the time interval [50-51 s]. From the visualised scanpaths it can be derived that during the first half of this interval, the scanpaths are still distributed rather evenly over the entire view. However, during the second part of the interaction, [50.50-51.00 s], the attention of the users is again drawn to the list with the five names. At this moment, the new map display is not in its final position but the list with five names has already changed. Again the search of the names on the map starts in its upper part, but now the scanpaths have more straight and diagonal directions. Two seconds after the end of the simulation, the search pattern of the users is again more evenly spread across the map’s content.
Figure 3.4: Time series to visualise the evolution of the scanpaths over time (in s)

The scanpaths visualised in Figure 3.5 are a result of the combination of two filtering parameters, namely two attributes: participant (rows) and demo-map (columns). This type of filtering allows detecting if there are any differences in the search behaviour of a user when a different map – and thus a different layout in labels – was presented, on the one hand, but also if there are differences in the search behaviour of different users on a same map. Person 1, for example, does not look very often in the list with names on the right, whereas person 2 and person 3 check this list more frequently. The search patterns of person 2 are very straight whereas these of person 1 appear to be more chaotic and rather long. The background of the maps (the distribution of the labels) also has an influence on the
scanpaths, but this is less clear than the personal differences mentioned before. Map 2 has for example a collection of scanpaths going from the list on the top right to the middle of the map on the left whereas on map 3 this collection is directed towards the lower right and left corner of the map.

Figure 3.5: Filtering the scanpaths based on two attributes: participant and demo-map

The scanpaths in the first picture of Figure 3.6 are only filtered on one attribute: the demo-map that is used. In the two other pictures in Figure 3.6, the movement data is also filtered on a time interval. The second picture shows all eye movements before the simulated pan-operation whereas in the last picture only the ones after the simulation are visualised. The blue polygons presented along with the scanpaths are the Area of Interests (AOIs): the location of the ten labels a user had to locate plus the location of the list with the five names. Some patterns in the movement data are showing when applying this last filtering strategy, but stricter simplification and/or aggregation of this subset of the data is needed to be able to retain only the main
structures in the trajectories. These simplification and aggregation techniques are described in more detail in Section 3.3.2.

Figure 3.6: Filtering the scanpaths based on one demo-map and time interval

3.3.2. **Simplification and aggregation of the data**

As indicated before, an eye tracking experiment quickly results in a vast amount of data from which no meaningful patterns of conclusions can be derived as such. The patterns are hidden in the somewhat chaotic representation of all scanpaths registered during the experiment. Similar trajectories may be present in the data, but they are concealed by other scanpaths crossing the same area. Aggregating and simplifying the trajectories of the scanpaths is an essential technique to retain only the relevant structures in the data and filter out the minor deviation from the main search strategy. The Visual Analytics Toolkit has a number of options which allow simplifying and aggregating the data with the aid of standard or user defined areas in the view.

One function to aggregate the data in The Visual Analytics Toolkit is called ‘Generalise and Summarise Trajectories’. With this function, the area is divided into a set of Voronoi polygons that reflect the density of PORs and minimises the distortion of the scanpaths. Based on the Voronoi tessellation, the trajectories are divided into individual segments which are subsequently aggregated if they go from and to the same Voronoi polygon pair. A detailed description of this function and its related algorithms can be found in Andrienko and Andrienko (2011).
After some testing it was decided that the following parameter values give the best results regarding the Voronoi tessellation and the clustering of the eye movements:
- Minimal angle of direction change (degree): 10
- Minimum radius around a position (pixels): 30
- Maximum radius around a position (pixels): 60

Different time filters can be applied to the clustered data to visualise only a small subset in an aggregated way. In Figure 3.7, this clustering technique is applied to the eye movement data of all users on one demo-map. The first picture shows the eye movements before the simulation and the second one the movements after the simulation. Clustering this dataset allows visualising the main movement patterns and thus facilitating its interpretation. In the last picture of Figure 3.7, the aggregated moves of the scanpaths during the first two seconds only are depicted. This result is in correspondence with the output from the time series discussed earlier in Section 3.3.1.

Furthermore, the data can be simplified and summarised based on a number of user defined areas. Regarding eye movement data, the AOIs can be used to investigate the number of scanpaths between these areas. Especially the number of movements between the list with names and the locations of the names is very interesting. These simplified trajectories which are visualised as moves between areas, can furthermore be aggregated in clusters. In this case, the clustering is not based on the calculated Voronoi tessellation, but on the user defined areas. These simplified and summarised moves between the Areas of Interest are depicted in Figure 3.8 and Figure 3.9 respectively, both for before (a.) and after (b.) the simulation, for all users but for one demo-map only.
Both pictures before the interaction (a.) in Figure 3.8 and Figure 3.9 show that the movements are directed towards the light blue rectangles. These rectangles correspond to the locations of the five names which appear in the list before the simulation. The darker blue rectangles correspond to the five labels which have to be located after the interaction. During the first interval, almost no moves are directed towards these latter rectangles. The pictures of the aggregated and simplified movements after the interaction (b.) show the opposite structure. In this case, almost all movements are directed towards the darker rectangle and thus to the locations of the five labels which had to be located after the simulation.
3.3.3. **Dividing the trajectories or scanpaths**

The Visual Analytics Toolkit also has the possibility to split up existing trajectories based on (user defined) areas they cross. This is a very interesting tool to filter out the part of the scanpaths between two AOIs, visualising the user’s search strategy. In this case, the trajectories between when the user looked in the list (which is defined as an AOI) and when he found a specific label (also indicated by an AOI) can be obtained. The resulting (partial) scanpaths for finding a specific label on a certain map are depicted in Figure 3.10. In each picture, the scanpaths for finding a specific label on a map for all users are depicted. The rectangles on the map correspond to the locations of the labels on the map which have to be found before (light blue) and after (dark blue rectangles) the simulation. These divided and selected trajectories indicate that some labels are more easily found by the participant than others. Some labels after the simulation are located directly with a straight line between the list and the label: Mia, Bokin, and Lankoe. These labels were already present in the list before the interaction (*) and the some users remembered its relative location.
3.4. Conclusion

The functions available in The Visual Analytics Toolkit include selection of the (eye) movement data based on attributes and/or time intervals, aggregation of the data based on a Voronoi tessellation or based on user defined areas, simplification of the data based on crossed (user defined) areas and subdivision of the trajectories (scanpaths) based on crossed areas. These techniques are applied to data in the main memory. However, the applicability of the approach is limited to about 20,000
– 50,000 scanpaths with up to 1,000,000 fixation points. Larger data sets require out-of-memory processing that integrates visual analytics with database sampling (Andrienko, et al., 2009).

The results obtained from the current study have given insights in how users process the information on a map when it is displayed for the first time and after an (simulated) interaction. Detecting patterns in the eye movement data is possible since the spatial dimension of the actual trajectories or scanpaths are visualised and analysed which is in contrast to the conventional quantitative approach in the analysis of the eye tracking data. Patterns in this movement data – such as the evolution of the search behaviour over time and the personal search strategies discussed above – cannot be detected when only analysing the measurements related to the fixations and saccades (e.g. duration and number).

References


Chapter 4

Investigating the Effectiveness of an Efficient Label Placement Method Using Eye Movement Data

Modified from:

This chapter focuses on improving the efficiency and effectiveness of dynamic and interactive maps in relation to the user. Two different label placement options are considered and implemented in user tests. We tested 30 participants while they were working on a dynamic and interactive map display. Their task was to locate geographical names on each of the presented maps. Their eye movements were registered together with the time at which a given label was found. The gathered data reveals no difference in the user’s response times, neither in the number and the duration of the fixations between both map designs. The results of this study show that the efficiency of label placement algorithms can be improved without disturbing the user’s cognitive map. Consequently, we created a more efficient map without affecting its effectiveness towards the user.

Keywords: label placement; interactive maps; eye tracking; user study
4.1. Introduction
The last two decades have seen a significant growth in the number of cartographic products available on the Internet. The terms neo-geography and neo-cartography have become common in the literature (Haklay, et al., 2008; Jobst & Döllner, 2008; Kraak & He, 2009; Turner, 2006). These terms are closely linked with the latest technical, multimedia, and Web developments, such as Web 2.0, which combine these developments with traditional GIS, cartography, and imagery research. This evolution has greatly facilitated the accessibility and availability of geographic data, resulting in an enormous increase in web map use.

However, these new possibilities and techniques also bring new challenges. For one there is the key issue of efficiency and effectiveness of web maps. Since these terms are used in different research and application fields – each using their own definitions – there might be some confusion regarding their exact meaning. In this chapter, the term efficiency, is used from an algorithmic point of view: how efficient is a certain algorithm or heuristic? One of the key factors used to measure this efficiency is the speed with which a certain algorithm can accomplish its goal (with a certain level of quality). In the case digital (Internet) maps, the construction of optimal algorithms and data structures to support or ‘calculate’ the visualisation is essential. The main challenge is thus related to the interaction possibilities, because a new visualisation must be calculated on the fly each time the user alters an element of the display (e.g., zoom factor, extent, colours, or layer visualisation). To illustrate, if a user zooms in on a map, the extent of the map that will be displayed after the interaction changes: a smaller region is displayed, but with greater detail. The algorithm must retrieve the objects within the limits of this new extent; only those objects that are linked with the correct level of detail or zoom factor are considered. The name of a country, for example, may not be displayed on a map if the user zooms in on a little village within that country. Furthermore, users do not want to wait several seconds before a new visualisation or display is generated each time they interact with it. Calculation times lasting more than one second are deemed unacceptable (Freeman, 2005). The efficiency of a web map is thus related to the fast calculation of a new display after an interaction.
The term **effectiveness** is in this chapter used from a human information processing point of view: how well can a certain person process the information presented to him? The effectiveness of a map is thus related to how well the user can process the visual information on it. Therefore, an effective map is one which can easily be processed by the user, although the information on it might be rather complex. Consequently, the effectiveness of a map is, on the one hand, inherently related to how the information is presented: the map design. On the one hand, it is also related to how users process this information. To be able to create these effective maps or map designs, it is crucial to understand how users store and process the visual information presented to them. This is in turn related to the users' cognitive processes or cognition (Harrower, 2007; Slocum, *et al*., 2001). Research on cognition considers different aspects of users their internal mental structures and processes related to, among others, perception, learning, memory, thinking, reasoning, and communication (Montello, 2002, 2009). As mentioned by Harrower (2007) and Slocum, *et al* (2001), the challenge of creating effective maps is even bigger in the case of the highly dynamic and interactive maps: the limits of the map readers’ cognitive processing capabilities need to be considered.

When combining both aspects, it can be concluded that an exceptionally efficient algorithm is useless if the result is chaotic and thus difficult to process by the user. In order to create effective maps, the algorithms need to consider a number of criteria related to how the information needs to be visualised correctly. These criteria are derived from the knowledge of the users’ cognitive processes and might have a negative effect on the algorithms' efficiency. As a consequence, a balance has to be found between the efficiency of the algorithm on the one hand and the effectiveness of the map (or the quality of the design) on the other hand, especially when dealing with highly interactive maps.

4.1.1. **Towards efficient cartographic displays**

A cartographic display consists out of a number of different symbol types, each contributing to the presentation of information. Name labels are one of the most important symbol types because they contribute significantly to orienting the user within the map. Map users prefer the name labels on a map because less cartographic knowledge is necessary to understand the map content. These labels
communicate information which cannot be presented with other types of symbols. Furthermore, labels (and their placement) can have a tremendous impact on the quality of a map because their presence informs the map user about much more than only the name of the geographic object associated to it (Barrault, 1998). They fulfil a number of essential functions related to the identification, location, orientation, and hierarchical structure of the features. They are thus essential to assisting the user in interpreting the content of the map correctly. Imhof (1975, p.128) also stated that “[...] good form and placing of type makes a good map” (Ahn & Freeman, 1984; Wood, 2000).

Both Cuenin (1972) and Imhof (1975) proposed a limited number of basic rules to place labels on maps to create effective results. Imhof (1975, p.129) summarises these rules as follows:

1. The names should be easily read, easily discriminated and easily and quickly located. […]
2. The name and the object to which it belongs should be easily recognised. […]
3. Names should disturb other map contents as little as possible. […]
4. Names should assist directly in revealing spatial situation, territorial extent, connections, importance, and differentiation of objects.
5. Type arrangement should reflect the classification and hierarchy of objects on the map. […]
6. Names should not be evenly dispersed over the map, nor should names be densely clustered. […]”

Although these basic principles were defined before the popularity of digital maps, they still hold true today for the dynamic and interactive maps on the Internet. Until recently, label placement on maps was done completely manually. This was a time consuming task, which could take up to 50% or more of the whole production time of a map (Yoeli, 1972). With the evolution of technological possibilities, many studies have been devoted to automating this process. Automatic label placement only takes the first three, apparently simple, rules of Imhof (1975) into account.

However, the label placement problem – even when only considering point objects – is an NP-hard problem, which means no optimal solution for the problem can be
found in an acceptable time span. Testing all possible placement options results in an exponential time complexity (Marks & Shieber, 1991). This time complexity can be on the order of $O(4^n)$ if each point object has four candidate positions where a label can be placed, with $n$ the number of points in the dataset. In 1996, the ACM Computational Geometry Task Force identified the label placement problem as an important research field in Discrete Computational Geometry (Chazelle, et al., 1996).

Consequently, suitable heuristics and meta-heuristics are needed to find a near-optimal solution for a certain situation, such as a specific object type. Over the years, different types of heuristic approaches have been developed, among others in the field of computational geometry. A number of these optimisation algorithms have been studied in the light of the label placement problem and have subsequently been adapted to improve their efficiency for this specific problem. Among the most interesting heuristics that deal with the calculation of label placement are: the Tabu Search Heuristic (Yamamoto, et al., 2002); Simulated Annealing (Christensen, et al., 1995; Zoraster, 1997); the Greedy Randomised Adaptive Search Procedure (Cravo, et al., 2008); Multiple Choice Integer Programming (Zoraster, 1990); and Genetic Algorithms (van Dijk, 2001). A number of solutions have also been proposed to specifically label linear objects (Edmondson, et al., 1996; Wolff, et al., 2000), and polygonal objects (Barrault, 2001). Furthermore, different applications for map displays have been taken into account in the construction of the heuristics. Automatic label placement for real-time displays has recently been given much attention (Been, et al., 2006; Mote, 2007; Yamamoto, et al., 2005; Zhang & Harrie, 2006).

Data structures are also indispensible elements, as it is important to store and process data efficiently. This especially holds true when large amounts of labels need to be considered. The use of conflict graphs (Kakoulis & Tollis, 2001; Wagner, et al., 2001) in a pre-processing phase allows structuring the data in a way which is surveyable, speeding up the actual calculations considerably. In this pre-processing phase, potential conflicts between the labels’ candidate positions are identified and stored, using a graph. The proposal of ‘sliding labels’ (Strijk & van Kreveld, 2002; van Kreveld, et al., 1999) is another example of a (data) structure that results in a
more flexible selection of candidate positions in such a way that more labels can be placed. However, despite the many extensive studies related to the label placement problem, no optimal solution has yet been found. This problem is also effectively summarised by Wood (2000, p.5): “[…] no one algorithm seems to be capable of recognizing the many considerations that a skilled human cartographer is capable of making in lettering a map.”

The studies discussed above mainly focus on finding efficient label placement algorithms (through the use of different heuristics and data structures), but the user side of the problem has been neglected. These heuristics are based on a number of rules concerning a good placement of the labels in relation to the user, but their final results have not been tested in an actual user study. This user side of the label placement problem is thus related to how users perceive the results of the (efficient) algorithms. A highly efficient algorithm is useless if the results it generates appear chaotic and difficult to interpret. Therefore, user studies are necessary to discover how users interpret the results from the heuristics.

van Dijk, et al. (2002) have developed a method of assessing the quality of the output of the heuristics. This method employs a quality function that measures how well a certain heuristic places labels on a map. The criteria used in this function are based on the six principles proposed by Imhof (1975) and therefore have a more theoretical background. Bradstreet, et al. (2005, p.1937) also described map labelling as a process in which labels should be arranged on the map “such that the result is clear enough for the map to be useful for a given user and task.” They also note that ‘clear enough’ means the labels are readable to the user and that it is easy to identify the label with its associated site and vice versa.

In both cases, no actual users were involved in these quality assessments, making them less suitable for investigating the effectiveness of label placement in respect to the user. To truly investigate the effectiveness of the map display, a user study has to be designed. A wide range of techniques are available for conducting user studies, but not all techniques are equally suitable for a given problem (Downs & Stea, 1977; Rubin & Chisnell, 2008; van Elzakker, 2004). This chapter focuses on improving the effectiveness of dynamic and interactive map displays towards the user. More
specifically, the influence of different label placement options on the effectiveness of a dynamic map display is considered. In the next section, a label placement algorithm with an improved efficiency is described. A user study is constructed to test whether the user experiences any influence due to the different label layouts created by the proposed algorithm.

4.1.2. Improved efficiency

As mentioned in the previous section, a wide range of label placement algorithms already exists, all using a slightly different method to improve their (algorithmic) efficiency. However, they all start with the same operation: determine the objects within the current extent of the map. The most optimal location of the labels associated with these objects has to be calculated by the algorithm. How this is done exactly depends on the specific algorithm used. The ‘current extent’ of a digital map corresponds to the part of the map that is visible on the screen. What is more, the current extent (or view) of highly interactive maps can change very often because users can zoom, pan, etc.

When considering panning, a lot of unnecessary calculations are done (by the algorithm) after each interaction. Panning is basically a translation of the map’s extent in a certain direction over a certain distance. Users often pan on a map to explore the surroundings of the current map. As a consequence, they might want to keep a rather large part of the map visible in order to orientate the newly displayed view in relation to the original one (from before the pan operation). We will call this part of the map that remains visible after the pan operation ‘the overlap zone’.

This type of interaction is illustrated in Figure 4.1a and Figure 4.1b. The large blue rectangles correspond to the current view (or extent) of the map; the small rectangles represent the labels on the map. Figure 4.1a depicts the initial view and Figure 4.1b the view after the interaction: a horizontal pan operation to the right. Current label placement algorithms (re)calculate an ‘optimal’ position for all labels within the new extent. These are the red labels in Figure 4.1b. Like most algorithms do, the level of preference related to a certain label position (relative to the associated point) is taken into account; the upper right position is the most preferred candidate position to place a label (Imhof, 1975). Compared to the initial
view (see Figure 4.1a), only the dark red labels’ relative positions have changed. Although all labels that are situated within the extent of the overlap zone are included in the calculations of the label placement algorithm, a more optimal position could be found for only a small number of them. As a consequence, a large number of unnecessary calculations have been done or a large number of labels have unnecessarily been included in the calculations.

![Figure 4.1: Illustration of the number of labels considered by the algorithms during panning](image)

Decreasing the number of labels considered by the label placement algorithm would significantly improve its efficiency. A closer look at the maps shows that most labels can remain in the same relative position, because they were already unambiguously positioned in the former view (before the interaction). Only the labels that were not visible in the former view and labels that could become unreadable after the interaction need to be included in the calculations. The latter group constitutes the labels located at the border of the current view, which could be clipped at the edges. Consequently, all labels that were visible in the former view and were not located near the border (and thus not in danger of being cut in half) can remain in the same relative position and should not be reconsidered by the label placement algorithm.
A rather large portion of the labels will fit this category, because the panning operation is often used to explore the surroundings of a certain location. As mentioned before, the user wants to keep a part of the former view visible after the interaction to facilitate the orientation of the newly displayed area.

The outcome when using the algorithm with the improved efficiency is depicted in Figure 4.1c. The initial view (before the panning) still corresponds to Figure 4.1a. The green labels in this third picture are the ones that could safely remain in the same relative position without the danger of rendering them unreadable. Only the labels in red should be (re)considered by the label placement algorithm. The red labels on the right side of the view are those which were not yet displayed in the initial view. Since they were not yet displayed in the initial view, the algorithm has to calculate a position for them, regardless which algorithm is used. The dark red labels, were already present in the former view, but are located near the border of the current view and are thus in danger of being clipped at the edge. In this case, this is only one label at the left side of the view.

The advantage of the algorithms used today is the higher level of quality of the final map. Every new view is recalculated, independently of the results of previous calculations. This means that for every view, the most optimal position for every label can be considered. However, this higher level of quality pays a price in terms of (algorithmic) efficiency because a lot of unnecessary calculations are done. The algorithm proposed in this chapter has a higher level of efficiency since fewer calculations have to be done. Labels which were already placed in the former view are not recalculated; except for a few that otherwise would be cut in half by the border of the map.

The disadvantage of this improved algorithm (in terms of efficiency) concerns the quality of the outcome. Fewer labels are placed in their optimal position, although it would be possible to place them there. Furthermore, subsequent panning operations over a small distance decrease the quality even more: with every interaction, fewer labels will be placed in the optimal position. This lower level of quality might affect the effectiveness of the map in a negative way. In the example of Figure 4.1a, a number of labels are placed in the upper left corner (relative to their point object) in
the initial view because otherwise they would be cut in half by the right border of the map. After the pan operation, they are still located in the upper left corner (see Figure 4.1c, which depicts the outcome of the proposed algorithm), although now there is space to place them in the most preferred position, which is the upper right corner (see Figure 4.1b, which depicts the outcome of existing algorithms). Other research concerning automatic label placement on interactive maps suggests that popping or moving of labels is also not desirable because it is distracting (Been, et al., 2006). It might be more difficult for a user to orientate the map if a number of labels are in a different relative position after the interaction, because this increases complexity of the map (Harrower, 2007).

The algorithm described above improves the efficiency of the label placement algorithm, but is the outcome also beneficial towards the users? The result of the algorithm also has consequences on how the labels are displayed on the map, especially after an interaction. From the elements described above, it can be derived that the outcome of the proposed algorithm has both negative and positive elements regarding the quality of the map and thus regarding its effectiveness: fewer labels are placed in a position with a higher level of preference, but the users are less distracted after the interaction because the relative position of more labels remains the same. The goal of the current study is to investigate if the application of this algorithm might have a negative influence on the user’s cognitive processes. Therefore, a suitable method, able to register changes in the user’s attentive behaviour and cognitive processes while working with a dynamic medium, needs to be selected.

In this chapter, we focus on the panning operations. As mentioned before, panning is an operation which is rather often executed independently from other operations – such as zooming, turning on layers, and changing font sizes – to explore the surroundings of the current map view. Improving the efficiency of the label placement algorithms related to this type of interaction would therefore improve the overall processing times when (re)calculating a map’s visualisation. Furthermore, turning on layers, changing font sizes, and other similar interaction possibilities are not as widely integrated in digital maps as zooming and panning. The proposed algorithm might also be beneficial when considering zooming, since some of the
labels can remain in the same relative position. However, zooming on a map also influences the priority levels of the labels on the map: some labels should not be placed at a certain zoom level, whereas others might appear on that level.

4.1.3. Tracking effectiveness

Nivala (2007) studied the familiarity of cartographers with usability engineering and its methods. Their results support the suitability of usability engineering for map application design, and they showed that the methods are slowly being incorporated in the design process. Consequently, research on the improvement of the effectiveness of map displays towards the user is almost non-existent, certainly when considering the influence of label placement options on a user’s cognitive processes. This lack of research is in sharp contrast with the amount of research devoted to enhancing the efficiency of label placement methods.

Duchowski (2007) and Rayner (1998) described the close link between a person’s eye movements and their cognitive processes. In the past, concerns about the eye tracking method have been expressed because there is no one-to-one relation between eye movements and attentive behaviour: a person can shift his attention without moving his eyes. However, with complex stimuli (or when complex cognitive processes are triggered, causing a high cognitive load), it is more efficient to move the eyes than to move attention. As a consequence, it can safely be assumed that there is a tight link between eye movements and attentive behaviour when a person has to solve a complex task (Duchowski, 2007; Jacob & Karn, 2003; Rayner, 1998).

During the last decade, the eye tracking method has been successfully applied to a wide range of studies concerning Human Computer Interaction (HCI) and usability research (Djamasbi, et al., 2010; Jacob & Karn, 2003; Joachims, et al., 2005; Salvucci & Anderson, 2001; Schiessl, et al., 2003; Zambarbieri, et al., 2008). Poole and Ball (2006, p.211) stated that this method “can help HCI researchers understand visual and display-based information processing and the factors that may impact upon the usability of system interfaces,” which was also previously described in the work of Henderson and Hollingworth (1998). Moreover, Fleetwood and Byrne (2006) studied and modelled the visual search of computer users on a graphical user interface based on eye tracking data. This study has some ground in common with
the experiments described in this chapter. However, the study by Fleetwood and Byrne (2006) focused on the design of icons in a graphical user interface, whereas the experiments described in this chapter are directed to the improvement of the quality of the map design for the user.

Furthermore, the eye tracking method has been introduced in the field of cartography to study, for example, the effectiveness of a new symbol design for a topographic map of Denmark (Brodersen, et al., 2001). These authors concluded that “the method as a whole is promising for future studies of map reading and map design as it reveals interesting information about a map’s design, the user’s behaviour, and the correlation between these” (Brodersen, et al., 2001, p.57). More recently, the eye tracking method has been integrated into studies that evaluated the effectiveness of interactive maps and small-multiple map displays (Çöltekin, et al., 2009; Fabrikant, et al., 2008).

4.2. Study design

In the present experiment, the effectiveness of different map designs – which are a result of the application of different label placement algorithms – towards the user was tested. It was chosen to use a controlled study design in order to avoid as many (unforeseen) influencing factors as possible and to be able to study one element in more detail. In this experiment only one type of interaction was thus considered: panning. This choice was based on the fact that panning is possible on almost every digital map and it is an operation which is commonly executed on a map.

Each map design included in study was linked to one of two specific label placement algorithms, which are in turn linked to the efficiency of the associated algorithms. This process allowed testing whether improvements in the efficiency of the label placement algorithms – with different visual outputs – had any (negative) influence on the user’s cognitive processes and thus on the effectiveness of the map display. The design of the study is described in more detail in the next sections.
4.2.1. Participants
Since the study has a within user design, a group of 30 students was selected. All participants speak Dutch, have a similar educational background, and belong to the same age group (between 18 and 20 year). The participants registered to the experiment on a voluntary basis. None of them received any previous cartographic training.

4.2.2. Task and stimuli
To test the different map designs, a number of map stimuli were constructed. An example of such a map stimulus is presented in Figure 4.2. Each of these map stimuli had the same simple background: three areas filled with a pastel colour. On this background, point objects and their associated labels are presented. These points and the names linked to them originate from real-world locations to obtain a realistic distribution. The 20 areas selected to construct the map stimuli had to meet a number of criteria. First, the participants had to be unfamiliar with the region. Second, the region could contain only names that are rather short. Third, it had to contain a constant number of names. Fourth, the region cannot contain any well known names.

Combining these 20 areas with the two label placement options (with and without the application of the proposed algorithm) resulted in a total of 40 map stimuli. Because the duration of a user study is preferably limited, these 40 maps were divided into two sets of 20. On ten maps of each set, the improved algorithm was applied: only labels near the border could be relocated after an interaction. On the remaining ten maps, all labels in the map image could be relocated after the interaction. This subdivision prevented a learning effect, as the same region will not be displayed twice to a single user. All of the maps were displayed in a random order during the study.

During the display of such a map stimulus, three different time intervals could be identified. The first interval was the initial view of the map, displayed to the user for 50 seconds. Next, a horizontal pan operation was simulated. This pan operation corresponded to a translation of the map image to the right. Finally, the view after the interaction was presented until the user located all names and ended the trial.
These three stages are depicted in Figure 4.2. The definition of the length of the initial interval was based on the results of an (unpublished) experiment, which we conducted prior to the actual experiment. The study design of this pilot study was the same as the one described in this chapter, except that only the reaction times were measured and the time of the first interval was fixed at 90 seconds. The results of this pilot study revealed that almost all participants could find the five labels within 50 seconds.

On the right side of each map stimulus a list of five names that were visible during the whole trial. The content of this list changed after the interaction: three names remained in the list and two were new ones. Due to the interaction, the absolute position of all labels on the map changed: they moved horizontally to the left over a certain fixed distance (the pan distance). In the list itself, the three names that were retained were also displayed on a different position after the interaction.

During the study, each participant had to repeat a number of tasks on each of the 20 map stimuli. First, he had to locate the five names in the initial view. Each time the participant found a name he had to push a button, resulting in a response time measurement. In order to avoid distracting the participant from the actual task, he had to push the same button for each label. After 50 seconds, the horizontal pan operation was simulated, during which the current view and the list with names changed. Next, the participant again had to locate these five names in the new view and push a button each time he found a name. When the participant had located all names, he could terminate the current trial himself and proceed to the next map.
The instructions were read out loud to each of the participants to make sure that everybody got them in exactly the same way. They contained a detailed description of the structure of the study, including how many names they had to find, the number of trials, duration of the initial view, the pan operation, and a description of the equipment. Furthermore, each user had to do two demo-trials before starting with the actual test. The goal of these demo-trials was to make the participant familiar with the task, stimuli, and equipment. As a consequence, nothing was recorded during the demo-trials: the participant could still make mistakes or ask additional questions.

The task described above corresponds to a visual search. This is a task that is rather often executed on a map: the user wants to locate multiple points of interest (POIs) on a map, such as tourist attractions or cities. Phillips, et al. (1978) also considered searching to be a good task to evaluate a map. They used the time to find a name as a measure to evaluate different types of typography used on a map. The disadvantage of working with a map on a relatively small screen is that these POIs are often not situated within the extent of one view (with a certain scale). The user has to execute a pan operation to get the objects within one view. Consequently, he has to relocate the POIs from the previous view and the POIs in the new view.

In order to obtain a similar effect in the controlled environment of the experiment, it was decided to list five names, from which three were repeated after the interaction. This also prevented the participant from tracking the relative position of the labels during the simulated pan operation. Since the names which remained in the list were chosen randomly, the user could not know beforehand which label he would have to relocate after the interaction. Otherwise, the results would not have been comparable between the different maps. In order to keep the participant’s search behaviour as natural as possible, he could choose whether he wanted to search for the five names in the given order or randomly. As can be derived from Figure 4.1, the difference between the two algorithms is situated in the overlap zone. As a consequence, after the interaction, the user has to relocate the majority of the names in this part of the map.
Figure 4.3 gives an overview of the different types of cognitive processes that could be identified during this visual search on the map. When the map stimulus is displayed, the participant will read one or more names from the list displayed on the right side of the map. Then he will look at the map to initiate the search. If the map is displayed for the first time, he will need some extra time to orientate it. During his search, he will try to identify potential label candidates: labels which look like the one he had read form the list earlier (e.g., length, shape of subsequent characters). When he has identified such a potential candidate, the participant will compare it with the information stored in his (working) memory. If he finds a match he can either return to the list with names to read one or more new ones or, if he still has one or more names in memory, he will start searching again on the map. If no match was found, he will continue searching on the map. A more detailed discussion about these cognitive processes and their relation to the Cognitive Load Theory can be found in Chapter 2.

Figure 4.3: Overview of the users’ cognitive processes during the visual search task

After completing the 20 trials, the user had to fill in a questionnaire that gathered background information, structured in several themes: general information, known regions, familiarity with web maps, and feedback. The measurements from the tests could be distorted by certain factors such as the familiarity of a participant with a
certain region displayed during the tests. The questionnaire is thus as a tool to eliminate biases and explain these possible distortions in the data.

4.2.3. Apparatus and data
The Eye Tracking laboratory of the Department of Experimental Psychology (Ghent University, Belgium) is equipped with an Eye Link 1000 device. This eye tracking device from SR Research (Mississauga, Ontario, Canada) can sample a person’s Point Of Regard at a rate of 1,000 Hz (or once every millisecond). The eye movements are recorded from one eye only. The software used to program the study and subsequently record the eye movements is the Experiment Builder (SR Research). Data Viewer (SR Research) is used to conduct the initial processing and visualisation of the obtained raw eye movement data. This software also allows identifying eye movement metrics related to fixation and saccades. The monitor on which the stimuli were presented had a size of 21 inch and a resolution of 1280 by 1024 pixels.

During the study, a number of different elements were recorded simultaneously, which have to be combined in a correct way. First, response times were measured related to the moment a participant pressed a button indicating that he had found a label in the current view. A comparison of these response time measurements between different map designs can then indicate which design allows a faster retrieval of information.

However, when using this method alone, there is absolutely no information on which labels are found or whether the participant found the correct labels. Furthermore, there is no indication of how easily the information on the map can be processed by the user. This type of information can typically be derived from the eye movement metrics related to the fixations. The duration of the fixation, for example, gives an indication of the time a user needs to process a certain view. However, eye movement data alone do not allow detecting if a label is found by a user and if it is the correct label. The user can ‘accidently’ overlook the label without actually identifying it as one of the labels from the list.
Only the combined data from the button actions and the eye movements allow identifying whether a label is found by the user and which label this is. The time measurements from the button actions thus indicate when a user finds a label; the location of the fixation at that moment identifies the label that was found. But caution is needed when combining measurements from different sources as it may cause synchronisation errors. Synchronicity is achieved during this study because all recordings (eye movements and response time measurements) were processed simultaneously by one software system.

To get insights in the map users’ cognitive processes during this task, data from multiple different sources were analysed, providing information about different aspects of a user’s cognitive processes and attentive behaviour. Response time measurements are a type of performance measurement which is very often used in the field of human-computer interaction to evaluate the design of, among others, user interfaces and web pages (Nielsen, 1993). In this case, these measurements indicate how fast a user can find a name on a map. A more effective map design would result in smaller response times because the user can process the (visual) information more easily (Harrower, 2007; Phillips, et al., 1978).

More detailed information about the participants’ cognitive processes can be derived from the eye movement metrics. The work of Jacob and Karn (2003) and Poole and Ball (2006) provide an excellent overview of a wide range of studies in which the eye movements of participants are analysed, which eye movement metrics are most often used, and what can be derived from them in terms of human-computer interaction. Since the goal of this research is to test the effectiveness of two different map designs (that are a result of two different label placement algorithms), the participants’ fixations are analysed: number of fixations and fixation duration. For example, the duration of the fixations may give insights in the difficulty with which a user processes the visual information presented to him. A very chaotic, and thus not very effective, map would consequently result in longer fixation durations (Jacob & Karn, 2003; Poole & Ball, 2006).

As mentioned before, a close link between a participants eye movements and his attentive behaviour can be assumed if complex cognitive processes are triggered
(corresponding to a high cognitive load). The task and the stimuli described above might seem rather simple, but they do trigger a high cognitive load: while the participant is searching for a label on the map, he is also trying to remember the locations of the names located previously. This results in an increasing cognitive load (see Chapter 2). Furthermore, the basic design of the maps and the controlled laboratory environment prevents the participants to get distracted, both due to internal (within the task or stimuli) or external (e.g., noises, striking elements in the surroundings) causes. Based on the considerations mentioned above, it can safely be assumed that the participants will not be distracted during the study and consequently, that there is a close link between the eye movements and his attentive behaviour. Hence, the metrics resulting from the eye moment study can be used to obtain insights in the underlying cognitive processes.

A qualitative, visual analysis of the eye movements also allows analysing the spatial dimension of the obtained data. Deviations or patterns in the users’ scanpaths, a succession of subsequent fixations and saccades, give insights in the users’ search behaviour and might indicate problems in the (map) design (see Chapter 3). Finally, these quantitative data sources are complemented with the results of a post study questionnaire. The goal of this questionnaire is to obtain information about the users’ preferences towards a certain type of label placement. A user’s preference towards a certain type of design might not always correspond to the most effective one (Nielsen, 1993).

4.3. Results and discussion
The results from two participants were removed from the dataset. The first participant was familiar with the presented regions and knew some of the names by heart. The other participant was extremely slow in locating the names. As a consequence, he could not execute the assignment correctly. From the remaining participants, all measurements related to locating a wrong name were removed from the dataset. Analyses of variance (ANOVAs) were performed across all participants with map design (border-labels relocated vs. all-labels relocated) as independent variables. The dependent variables were the mean reaction times across trials, fixation count, and fixation duration.
4.3.1. Relative response times

The results showed no difference between the border-labels relocated and all-labels relocated map design conditions ($F = .069, P = .793$). The mean reaction time for locating the ten labels was $5.614 \text{ s} (SD = 1.716 \text{ s})$ for the condition in which only labels at the border were relocated after an interaction and $5.570 \text{ s} (SD = 1.630 \text{ s})$ for the condition in which all labels could be relocated after the interaction. The mean values (in seconds) for locating the different labels in the two map designs are illustrated in Figure 4.4. In the map design border only labels near the border could be relocated after the interaction; in map design total all label could be relocated. Five labels were located before the simulated interaction (Before) and five after this interaction (After). Paired comparisons of the intervals of locating each label in both map designs (e.g., the first label that was found in the border-design vs. the total-design, the second label that was found in the border-design vs. the total-design) showed no significant differences for any of the intervals ($Ps > .106$).

![Figure 4.4: Mean response time measurements (with 95% confidence interval) for both types of label placement related to finding the five labels, before (a) and after (b) the interaction](image)

When a map is displayed to the user for the first time he has to orientate it. The time needed to do this is included in the time interval to locate the first label. Since no difference was found in this first time interval, it can be concluded that there was no difference in how fast the user can orientate the map. After the simulated pan operation, the user will orientate the map again: he wants to link the initial view with the map view after the interaction. Again, the orientation time is included in the time interval needed to locate the first label. Also in this case, no significant
difference could be detected between the two map designs (or label placement methods). As mentioned previously, users might be distracted after the simulated pan operation if some labels’ relative position suddenly changes after the interaction. However, in the case of the proposed label placement algorithm, no significant improvement could be detected. Another effect of the proposed algorithm on the map design is the lower quality in the final label placement: fewer labels are placed in their most preferred position. Since no significant difference could be detected in the corresponding response times of the two types of map design, it can be concluded that the users were not negatively affected in their visual search due to this lower quality.

4.3.2. Fixation count and duration
The fixation count corresponds to the number of fixations during the trial. Since this is inherently related to the length of the trial, a more objective metric is used: number of fixations per second. The number of fixations per second is closely linked with a second metric often used in eye movement research: fixation duration. Although both measurements are linked with each other, it might be very useful to consider them both in order to obtain a better understanding in the users’ cognitive processes. This duration of a fixation may give an indication about how the user processes the (visual) information presented to him. Longer fixation durations may indicate that the user is confused by or has difficulty in reading the display, but it could also indicate that the user find that part of the display interesting (Poole & Ball, 2006). However, the basic structure of the map stimuli in the experiment assures that all regions on the maps look equally interesting. Therefore, the latter explanation is not considered in this study. If the labels are placed in a very chaotic way on the map it is expected that the user needs more time to process the visual stimuli, resulting in longer fixation durations. Some of the labels on a map with a border-design, are placed in a less preferred position than with a total-design. Therefore, it might be expected that the proposed algorithm creates maps that are more difficult to process, resulting in deviating eye movement metrics such as longer fixation durations.

The mean number of fixation per second for both map designs is 3.843 fix/s ($SD = .436$ fix/s) with a minimum and maximum of respectively 2.425 fix/s and 5.036 fix/s.
The mean number of fix/s counts 3.829 for the border-map design ($SD = .451$ fix/s) and 3.856 ($SD = .422$ fix/s) for the total-map design. A one way ANOVA over all response times and with map design as a factor shows no significant difference on the level of the fixed factor ($F = .822; P = .365 > .05$). Since the map design type is related to how the labels are placed after an interaction, the label's positions might have a more pronounced influence on the user after the interaction. As a consequence, the dataset is further divided into two subsets: before and after the interaction. The mean fixation counts, related to both map designs, are illustrated in Figure 4.5. In this case, a larger difference with a border-map design can be noticed compared between before and after the interaction. However, a paired comparison between both map designs shows no significant difference in the results, both before ($F = .118, P = .731$) and after ($F = .787, P = .376$).

![Figure 4.5: Mean fixation count (fix/s, with 95% confidence interval) for both types of label placement, before and after the interaction](image)

The mean duration of fixations is .228 s ($SD = .032$ s) for both maps designs, with a minimum of .164 s and a maximum of .357 s. When only considering the border-map design this mean duration is .230 s ($SD = .342$ s), whereas this is .227 s ($SD = .032$ s) for the total-map design. Also on this dataset, a one way ANOVA is applied which results correspond to these from the fixation count: no significant difference on the level of the independent factor ($F = 1.439; P = .231 > .05$). The mean fixation durations, related to if they occurred before or after the interaction, are presented in Figure 4.6. Also in this case, the paired comparison showed no significant difference between the two map designs, both before ($F = .304, P = .582$) and after ($F = 1.228, P = .268$).
The eye movement metrics thus show no significant deviations between the application of the two different label placement methods. This means that, although fewer labels are placed in a more preferred position by the proposed algorithm, the users do not have more difficulty to process the information. This may be explained by the fact that it only concerns a limited number of labels that are not placed in a more optimal position: certain labels that were located near the border of the previous view and were therefore restricted in their placement.

4.3.3. Qualitative scanpath analysis
Besides the statistical analysis of the eye movement metrics, interesting elements regarding the users search behaviour might be obtained from a visual analysis of the scanpaths. These scanpaths are a succession of subsequent fixations and saccades. The spatial distribution of these scanpaths was analysed and compared using the Visual Analytics Toolkit, whose functions – and their applicability on eye movement data – are described in previous chapter (see Chapter 4). The scanpaths were filtered based on their attributes, using a fixed time interval. The distribution of the scanpaths related to the same area, but with the application of a different label placement algorithm, could thus be compared.

In order to create more comparable datasets, the eye movements of a certain map were filtered based on a fixed time interval. This is necessary because the users could end the trial themselves when they had located all labels. As a consequence, each trial had a different length. Since we had already fixed the time interval before
the interaction at 50 s, the same will be done to visualise the scanpaths after the interaction.

A comparison of the scanpaths of the corresponding map designs, both before and after the interaction, might reveals deviations or differences in how the users search on the map. For every combination, the scanpaths of the users were very similar. As an example, these scanpaths are illustrated for one of the maps in Figure 4.7. The two upper pictures correspond to maps which are created with a traditional labels placement algorithm (total-map design), whereas the two lower pictures are created with the proposed algorithm (border-map design). The pictures on the left side illustrate the scanpaths before the interaction and the pictures on the right side the ones after the interaction (during 50 s).

Figure 4.7: The users’ scanpaths when searching on one of the maps (top: total-map design, bottom: border-map design; left: before the interaction; right: after the interaction)
4.3.4. **Questionnaire**
A number of questions in the post study questionnaire were related to the two map designs presented to each participant. The goal of these questions was to discover whether the preference of the users towards one label placement type corresponds to the results of the statistical comparisons: time measurements and eye movement metrics.

These questions were formulated as follows (translated from Dutch):

> During the study, two types of maps were presented to you.
> - Describe the difference between both map types.
> - Which of the two types do you prefer the most?
> - Why do you prefer this type more than the other one?

All of the participants answered to the first question that they did not notice any difference between the 20 maps presented to them. Consequently, no useful answers could be derived from the two subsequent questions. But the answers on the first question clearly indicate that the participants did not notice the difference between the maps (or label layouts) presented to them.

4.4. **Conclusion and future work**
This chapter describes a study in which the effectiveness of an improved label placement algorithm (in terms of efficiency) towards the user is investigated. The combination of response time measurements and eye movement metrics allows identifying when a label was found and which label was found by the user. The gathered eye movement metrics, duration and number of fixations, allow detecting whether the different label layouts have an influence on the user’s cognitive processes and thus on its effectiveness towards the user.

The analysis of the participant’s response times shows that the proposed label placement algorithm does not have any significant influence on the user’s cognitive processes. The assignment can be completed in the same time interval on both map designs, which means that the different label layout does not affect the participant his orientation process, nor the his visual search on the map. More detailed analyses
of the user’s cognitive processes, using the eye movement metrics fixation count and fixation duration, confirm this finding even further. Both the number of fixations per seconds and the duration of these fixations are not significantly different. Furthermore, no deviations in the users' search behaviour could be derived from the visual analysis of the scanpaths. The questionnaire also revealed that the participants did not notice the difference between both label placement methods. It can thus be concluded that the two map designs, with or without the application of the proposed algorithm, are as difficult to interpret by the participants. Although both map designs created different visual outputs, the participants stated that they did not notice any difference between the maps presented to them. This confirms that the application of the more efficient algorithm does not influence the user in any way. The label placement algorithm is thus more efficient without affecting the user and thus the effectiveness of the map in any way.

However, these results were obtained during a controlled study, using a basic map design, a simple task and in a laboratory environment, which is different from a real life situation. So the question arises: how generalisable are these results? During the controlled study, the goal was to direct the focus of the participants only on the labels on the maps, avoiding other influencing or distracting factors (such as other map elements) as much as possible. As a consequence, if there would have been any effect of this adapted label placement algorithm on the map user, it would be expected to be the clearest in this study. More complex maps would lessen the effect of the lower qualitative label placement, as there are too many other elements on the map to which the focus of the user will be directed. Nevertheless, it should also be noted that the actual labels will be placed differently due to the other map elements, which may also affect the quality of the placement and thus its efficiency towards the user. In order to be able to answer this question objectively, a series of follow-up studies have to be conducted in which more complex map stimuli are used. Gradually adding more map elements to the stimuli would allow investigating which elements (or what level of complexity) would have to greatest impact on the map user.

In order to be able to create more effective maps, more information is needed on how users actually interpret maps: how do they orientate a map, how do they
process, store, and retrieve the visual information presented to them. Detailed insights in the user's cognitive structures are thus essential to understand these processes, especially while working with highly interactive and dynamic maps. Eye tracking is an excellent method to get in touch with the user's cognitive processes while working with visual stimuli. More detailed studies are planned to investigate how users orientate these dynamic maps, how their cognitive maps are created and what influences them.

References


Chapter 5

Understanding Expert and Novice Map Users: Reading and Interpretation

Modified from:
Ooms, K., De Maeyer, P., & Fack, V., Understanding expert and novice map users: Reading and interpretation. (Submitted to Cartography and Geographic Information Science)

The aim of this chapter is to gain a better understanding in how map users read and interpret the visual stimuli presented to them and how this can be influenced. Especially the difference between expert and novice map users is considered in this.

In a user study, with a between user design, the participants had to study four different screen maps, which were also manipulated to introduce deviations. The eye movements of expert and novice participants, 24 in total, were studied during these trials. Based on a grid of Areas of Interest, these eye movement recordings are analysed both visually and statistically. These, at least two dimensional, visual analyses are essential to study the spatial dimension of maps to locate problems in its design. The proposed analyses included the visualisation of eye movement metrics (fixation count and duration) in a 2D and 3D grid, including a statistical comparison of these grid cells. The results show that the users’ eye movements clearly reflect the main structuring elements on the map. The interpretation process of both user groups is also influenced by deviating colour use on the map as their attention is drawn to it. Furthermore, both user groups encounter difficulties when trying to interpret and store familiar map objects, which are not exactly the same as the ones they stored previously (caused by a mirror operation on the objects). These influences due to deviations in the map are found to be more pronounced with the novices than with the experts. Insights in how different types of map users read and interpret the map content are essential in this fast evolving era of digital cartographic products.

Keywords: user study; eye movements; cognitive cartography
5.1. Introduction

Due to the ever progressive advancements in technology, cartography has undergone a tremendous evolution during the last few decades. Already at the beginning of the new century, it was estimated that the number of maps distributed through the Internet on a daily basis exceeded the number of paper maps printed each day (Peterson, 2003). What is more, the Internet has made maps and cartography much more accessible to the general public. However, the main goal of these ‘modern’ cartographic products remains the same: communication.

Communication can be seen as a process in which different steps are involved, and as a consequence, cartography includes more than mapmaking alone. Several models of cartographic communication have been proposed since the 1960s. In its simplest form, the cartographic communication process can be seen as a transmission of source information (the world around us) to a recipient (the map reader). Maps visually represent the (spatial) information surrounding us, and communicate this to the users by applying some kind of code, the cartographic syntax. They are the channels that allow (and should facilitate) this transmission. Consequently, if a map’s design is not optimal, it could introduce noise in the communication process (Montello, 2002). A more elaborate model presented by Kolácný (1969) was most influential on cartography and was considered an essential aid in the studies regarding how easily maps can be interpreted (MacEachren, 1995).

The new technologies have a profound impact on how the cartographic products are displayed to the user, and thus on how the information is perceived and interpreted. Screen maps make way for new possibilities, such as animations and user interactions, but are also inherently linked to some critical limitations: resolution, size, colour use, etc. (Peterson, 2003). Furthermore, novice map users can, nowadays fairly easily, create cartographic products on the Internet, which are in turn used by expert or novice users.

Recently, some concerns regarding these evolutions in cartography (and GIScience) have risen. How effective are these new map displays? What effect do they have on the users’ cognitive processes? What are the limits of the map reader’s visual and cognitive processing capabilities? Several authors expressed these concerns and
concluded that more research is necessary regarding the cognitive issues in cartography and geographic information visualisation in general (e.g. Fabrikant & Lobben, 2009; Harrower, 2007; Montello, 2002, 2009; Slocum, et al., 2001).

Cartwright (2012) listed a number of issues related to cartographic products created by amateur map makers, among others regarding their design, maintenance, privacy, and data protection. In order to maintain an effective and efficient communication process, experts could introduce limitations or set guidelines for these amateur maps. As a consequence, it is essential to identify and study the differences between both user groups. In the next sections, the theoretical background regarding the cognitive structures and processes, necessary to interpret visual information (such as maps), is described.

5.1.1. How can we process visual stimuli?
Montello (2002) stated that cognitive cartography includes the study of knowledge structures and processes of map reading, such as perception, learning, and memory, to understand mapping. Taking into account the cartographic communication model, the first step in map use is the interpretation of the visual encoded information on these maps. This interpretation process consists in turn out of a number of subsequent steps (cognitive processes), which are linked to the structure of the human memory (Kulhavy & Stock, 1996). Atkinson and Shiffrin (1968) divided the memory in three structural components: the sensory memory, the short term memory, and the long term memory. This model had a major impact on the early studies in cognitive psychology, and is often called the modal model. Over time, modifications to this model have been proposed, among others by Anderson (1983), but its main structure is still accepted today.

The sensory memory records input from each of our senses, including vision, but this is quickly forgotten (less than two seconds). Some of the information in the sensory memory will be transferred to the short term memory, which also has a limited capacity. Nowadays, the short term memory is called the working memory (WM), and this term will also be used in the remainder of the chapter. In order to transfer the information from the WM to the long term memory (LTM) it has to be
rehearsed and thus learned (Matlin, 2002). The capacity of the LTM is considered virtually infinite.

A major contribution to the assessment of the capacity of the WM was made by Miller (1956), who found a (magical) number seven. In his article he described that the WM can hold seven chunks of information (give or take two), which poses a severe limit on the memory span in the WM. However, he also found that these chunks can vary in size. Information can be grouped together if it is closely linked, to form one chunk, which takes only one ‘slot’ in the WM. For example, the characters ‘m’, ‘a’, and ‘p’ will only form one chunk of information instead of three because they can be grouped into a meaningful word, ‘map’.

More recent studies re-examined Miller’s results and found that the WM’s capacity limit is actually less than seven. Most authors conclude than humans can hold, on average, four chunks of information in the WM (e.g. Cowan, 2001). Nevertheless, there is still some debate whether there is also a time limit on the information that can be held in the WM without rehearsal. Atkinson and Shiffrin (1968) stated, for example, that information stored in the WM would decay completely over time. Without rehearsal, this information is lost within 30 seconds. Nevertheless, the time limit of the WM is still not generally accepted and found controversial (Cowan, 2001).

In order to be able to explain the processes that take place in the WM, Baddely (1999) proposed a WM structure that consists out of three separate components: the phonological loop (which stores sounds), the visuo-spatial sketch pad (which stores visual and spatial information), and the central executive (which processes the information). This structure is depicted in Figure 5.1. Consequently, the WM is much more than a database that can store a certain number of chunks containing information. The information contained in the WM can constitute of information obtained through the sensory memory, but ‘old’ information retrieved from the LTM can also be held in the WM. The central executive enables the manipulation of these different information sources (Matlin, 2002). This memory structure is essential to understand how humans can process visual stimuli, and thus subsequently interpret them.
5.1.2. How do we interpret maps?
In order to interpret maps, and thus the visual information depicted on them, the map reader will use previous knowledge to process the stimuli registered by his (visual) senses. Two important cognitive processes form the basis of the interpretation process: attention and object recognition (Matlin, 2002).

Attention can be defined as a concentration of mental activity. To interpret a certain object, the user has to focus its attention on it. The subsequent foci of attention are (mostly unconsciously) guided by input from the users’ cognitive processes. Wolfe (1994) described a model (or set of rules) that predict where a user will focus his attention during a visual search. Recent findings in the field of change blindness also stressed the importance of attention in order to be able to detect changes in a visual scene (such as animations on maps) (among others, Rensink, 2002; Simons & Ambinder, 2005).

The next phase in the interpretation process is object recognition. The (often complex) input of the sensory stimulus is compared to other memories and then identified. This structure corresponds to Marr’s (1982) representational framework to vision. First, the visual sensory input is transformed into the primal sketch. This primal sketch is a (rather abstract) organisation of the retinal image in an array of cells that can indicate perceptual units (such as edges). Second, a 2.5D sketch is constructed which is a higher level representation of the array of cells, containing properties of the visible surfaces. In this stage, no integration or comparison with other knowledge (possibly from the LTM) has taken place. Finally, the user constructs a hierarchical 3D model of the perceived object, which contains the three dimensional structure and organisation of the object. Other frameworks have been
proposed, but Marr’s approach had a significant influence on the understanding of vision and information processing, because it is based on a computational theory (MacEachren, 1995).

In order to understand how users perceptually organise visual scenes, another influential approach has to be explained at this point: the Gestalt Theory. One of the ideas behind this theory is that humans will (unconsciously) try to organise or group what they see. Furthermore, the whole scene is greater than the sum of the parts. The Gestalt psychologist defined a number of rules (or laws) that humans apply during the organisation of a scene, such as the law of proximity, similarity, closure, experience, etc. Consequently, in order to interpret a visual scene, it is not sufficient to independently recognise the separate parts that are present in this scene (MacEachren, 1995; Matlin, 2002). This has a major impact on how cartographic stimuli are perceived and interpreted by the map users. Especially their level of expertise and background knowledge could have an influence on how the map users organise (group) the elements that constitute the map image and, consequently, how the whole map image is seen.

Several authors have also described different levels at which object recognition can take place. For example, Gerber (1981) identified three levels of successful object recognition. At the first level, the perception-recipe or pictorial level, the user can recognise the shape of an object that he has seen previously. At a higher level, the user also knows the name of the object; hence this called the label or pictorial-verbal level. Finally, if the user possesses other knowledge regarding the object, object recognition takes place at an even higher level: other knowledge about or verbal level. These levels are subsequently processed during the interpretation process of a certain object. When considering the interpretation or recognition of symbols on maps, Olson (1976) also identified three levels of processing: (1) compare symbol pairs, (2) recognise groups of symbols, (3) use the symbols to retrieve information from the map.

These different levels of object recognition are influenced by bottom-up or top-down processing of information. Top-down processing of is based on information that is already stored in the users’ memory (background information), which is
influenced by expertise. Bottom-up processing is solely based on the visual input, without linking it to previous knowledge (MacEachren, 1995). Hegarty, et al. (2010), for example, examined the influence of salience (bottom-up) and domain knowledge (top down) on map comprehension (in cased of weather map displays). They discovered that eye movements are mainly guided by top-down factors and that a good design facilitated the processing of task relevant visual features.

This interpretation process ‘consumes’ a part of the limited capacity of the WM. Another part of the WM’s capacity will be used to transfer the processed information to the LTM: the learning process. Sweller (1988), Bunch and Lloyd (2006), and Harrower (2007) give an excellent description of the ‘consumption’ of the limited capacity of the WM to process geographic and cartographic information. Their explanations are based on the *cognitive load theory*. This theory states that the WM is also limited in its processing capabilities; the amount of cognitive load it can take. Furthermore, different types of cognitive load can be identified, all contributing to the total amount of cognitive load. First is the *intrinsic cognitive load*, which is related to the complexity of the information on the map. Second, the *extraneous cognitive load* is influenced by the design of the map: poor design and distractions make it more difficult to interpret the map content and cause a rise of the cognitive load. Finally, the *germane cognitive load* is closely linked to the learning process. Consequently, maps containing overly complex information, represented in a chaotic way, will be very difficult to interpret by the user due to the high intrinsic and extraneous cognitive load. These types of maps cannot be interpreted or learned from in an efficient way, because there is no room left to address the germane cognitive load that facilitates the learning process.

In Chapter 2, the reaction time measurements and eye movements of expert and novice map users who performed a visual search on screen maps with a very basic design were investigated. Significant differences were identified between both user groups: experts seem to be able to interpret the map content much more efficiently than the novice users. The cognitive load theory was used as a basis to explain these findings. Alvarez and Cavenagh (2004) investigated the *visual information load*, which is related to the amount of detail of the perceived objects. A higher level of detail results in a slower processing rate for each object.
The maps presented during the user study described in Chapter 2 had a very simple design, containing no complex objects: only points with labels and three polygons visualised in a pastel colour. Consequently, these results cannot be generalised to a wide range of maps. The aim of this chapter is to extend this study with the incorporation of more complex (topographic) maps to gain even deeper insights in how different types of users process this visual information. Both amateur and experienced map makers can nowadays create and distribute cartographic products on the Internet, which are in turn accessible to a wide range of users (both experts and novices in cartography). As a consequence, it is important to consider both user groups in the experiments described in this chapter.

5.1.3. How to study ‘map interpretation’?

Eye tracking is a ‘direct’ method to study users’ cognitive processes. The participants in the user study do not have to reflect on their thoughts, using introspection or retrospection. Insights in their attentive behaviour are thus obtained without any user interference. Participants reflecting on their own thoughts results in subjective and unreliable results, often because they do not exactly know how their thoughts were formed: it is an automatic process (Nielsen, 1993; Rubin & Chisnell, 2008).

Using eye tracking, the position where a user is looking (Point of Regard, POR) is registered at a certain sampling rate. This gives insights in the user’s attentive behaviour: the location on which the focus of attention is directed at that moment in time. As mentioned before, attention is an essential step in object recognition and thus to interpret the whole map content. Besides the position of the POR, other usable metrics (such as fixation duration) can be derived from the eye movement data, which provide insights in the user’s cognitive processes during the interpretation of the visual content. Based on previous research regarding the eye tracking method, it can safely be assumed that a close link exists between these cognitive processes and the eye movement metrics (Duchowski, 2007; Jacob & Karn, 2003; Poole & Ball, 2006).

Only recently, a renewed interest in the use of eye tracking in cartographic studies is noticed (Brodersen, et al., 2001; Çöltekin, et al., 2009; Fabrikant, et al., 2008). This renewed interest is closely linked to recent need to gain a better understanding of
the cognitive processes (and limits) of map users while working with highly dynamic, interactive, animated screen maps. These insight are essential to be able to key the visualisation of future maps to the cognitive structures of the map users and in the end create more effective maps.

Eye movement data can be analysed in a number of different ways, from which two main categories can be identified: quantitative and qualitative methods. Quantitative methods use statistics (e.g. ANOVA) to identify significant differences between two categories: differences in map design (within user design) (see Chapter 4) or differences in user characteristics (between user design) (see Chapter 2). These types of analyses have a high level of objectivity, but often lack the spatial dimension. However, this dimension should not be ignored when studying maps and their design. A visual analytic approach to study eye movement data can handle this spatial dimension, as the distribution of the eye movements is considered. The visualisation of the participants’ scanpaths, for example, allows detecting patterns (see Chapter 3). However, caution has to be taken when interpreting the visual data to avoid subjective conclusions. Nevertheless, these different techniques shed light on other aspects of the eye movement data and thus on the related cognitive processes. In order to obtain the most elaborate overview of the cognitive processes that take place during the interpretation process of maps, it is good practice to combine different techniques.

5.2. Study design

5.2.1. Participants
Two groups of participants were selected to take part in the study, containing each 12 persons, with an equal share of males and females. The first group consisted out of experts in map use and cartography. All participants in this group had at least a Master degree in Geography or Geomatics and received cartographic training during their studies. Furthermore, at the time of the study they were employed at the Department of Geography at Ghent University. The second group consisted out of participants who did not receive any previous cartographic training and did not work with maps on a professional level. The average age of the participants was 23.8
years, with 25.9 years for the expert group and 21.4 years for the novice group. All participants took part in the study on a voluntary basis.

5.2.2. Tasks
The instructions were read out loud to each participant in order to avoid differences in how the task was interpreted due to different use of wording. Furthermore, at the start of the test, the participant could read through the instructions again on the screen. At this point, he could ask any questions if the instructions were not clear.

The instructions for the task were simple, but clear. The participant was told that a map would be shown on a screen and he had to remember the general structure of this map. He did not have to remember all details on the map (such as individual houses), but certainly the general structures, such as roads, rivers, forests, etc. In order to avoid any biases due to (time) pressure, the participant could execute this task at his own speed. If he found that he had studied this map long enough, he could remove the map display himself by pushing a button. The user was informed that the map was displayed for a maximum time of ten minutes, which should be more than enough to perform the task. This limit was introduced to keep the study manageable. It should be noted that this limit was reached only a few times over all participants and maps. During this task, the participants’ eye movements were registered.

In order to force the participants to interpret the map content, they had to execute a second task. Before the start of the study, the participants were also instructed that they would have to draw the map they had just seen, using paper and pencil. Likewise, no time limit was set on the drawing task; the participant just had to indicate when he was ready (see Chapter 6 for more information). This process of ‘remembering the map – drawing the map’ was repeated four times. After the completion of the fourth trial, the participant was asked to fill in a questionnaire. This post study questionnaire is used to obtain personal characteristics (expertise, age, gender, etc.), to verify potential familiarity of the participant with the presented regions, and to receive feedback.
5.2.3. **Stimuli**

During the user study, four different maps were presented on a screen. These stimuli were selected out of the Belgian topographic map series on 1 : 10 000. The maps were chosen as such that the map image would not be crowded with information; some obvious structures were visible (roads, rivers, forests, etc.), and the region was not well known. Familiarity with a certain area influences the user’s interpretation process caused by previous knowledge and should thus be avoided.

The four stimuli were depicted in the same order to all participants. The maps used during the study are depicted in Figure 5.2. This figure shows five different maps, although only four maps were presented to each user. This is due to a variation that was introduced in the third stimulus, which was only shown to half of the participants. Six experts and six novices saw the third map in its normal orientation; the other half of the test persons saw the map mirrored over its vertical central axis. This allows detecting whether the users’ scanpaths – which result from the interpretation process of this map – would also be mirrored.

Taking a closer look at these stimuli also shows that map 4 is a mirrored version of map 1; this time over the horizontal central axis. This means that each participant both saw the original map and the mirrored version, separated by two other stimuli (map 2 and map 3a or map 3b). This would provide insights in how the familiarity of the map image, which is also significantly different, influences the interpretation process of the map users. Finally, the second topographic map is characterised by a deviating use of colours to depict water bodies and the village backgrounds. The hue of both original colours (cyan and light yellow) has been changed over 180° into a light orange and purple respectively. Likewise to the familiarity of a region, the users might be familiar with the colour use of the Belgian topographic map (at a scale of 1 : 10 000). Deviations in this colour use could distract or confuse the users and thus influence the interpretation process. Furthermore, the differences between the experienced and novice map users in this will be considered in the analyses.
Figure 5.2: Stimuli depicted during the trials of the user study

5.2.4. **Apparatus and recordings**

The participants’ eye movements were registered with an EyeLink1000 eye tracking device from SR Research (Mississauga, Ontario, Canada). This device is installed in the eye tracking laboratory of the Department of Experimental Psychology (Ghent University). This desk mounted device with a chin rest can sample a user’s POR at a rate of 1,000 Hz. The maps were presented on a 21 inch monitor.

The software DataViewer (SR Research) was used to aggregate the raw data into meaningful measurements, such as fixations and saccades. Fixations correspond to time intervals at which the POR is relatively stable. At this moment the user is interpreting the visual information, and as a consequence, the metrics related to this
measurement are of most interest. The DataViewer has tools to create reports listing the number of fixations and average duration of these fixations for each trial. Furthermore, detailed information regarding these fixation metrics can be obtained separately for indicated Areas of Interest (AOIs). These AOIs are regions (squares in this case) that are compared with the eye movement data. For each AOI the number of fixation and the total duration of the fixations within its boundaries are listed (separately for each trial and participant).

5.3. Methodology and results
5.3.1. Statistical comparison: experts vs. novices
In the DataViewer software, a trial report can be exported. This report aggregates the available data (including eye movement measurements) per trial. The eye movement metrics of interest are the average duration of the participants’ fixations and the number of fixations per second. The average duration of the fixations can indicate the difficulty with which the visual stimulus is processed. Complex or chaotic stimuli, which are difficult to process by the user, due to a rise in the cognitive load, typically result in longer fixation durations. If a user would find a part of the visual stimulus particularly interesting, the duration of the fixations would also increase. The number of fixations a user can have per second is closely linked to the average fixation duration of that user. Longer fixation durations result in fewer fixations per seconds. However, studying both metrics can be useful to explain the results.

Table 5.1 lists the mean values (M) and standard deviations (SD) for the average fixation durations, the number of fixations per second and the duration of the trial, both for the expert and novice users. The last column in this table shows the results of a one way ANOVA between the two user groups. From this table it can be derived that the expert users have significantly shorter fixations than the novice users. Furthermore, experts can have significantly more fixations per second. These findings are in correspondence to the results described in Chapter 2, where the eye movement metrics resulting from a visual search on a very basic map design were described. The results obtained in the present study indicate that the findings from Chapter 2 can be generalised to a wider range of maps and applications.
Table 5.1: Statistical comparison (average fixation duration; number of fixations per second; duration of the trial) between expert and novice map users

<table>
<thead>
<tr>
<th></th>
<th>Experts</th>
<th></th>
<th>Novices</th>
<th></th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>F</td>
</tr>
<tr>
<td>AvgFixDur(s)</td>
<td>.308</td>
<td>.046</td>
<td>.339</td>
<td>.062</td>
<td>7.578</td>
</tr>
<tr>
<td>Fix/s</td>
<td>2.804</td>
<td>.467</td>
<td>2.566</td>
<td>.469</td>
<td>6.235</td>
</tr>
<tr>
<td>TrialDur (s)</td>
<td>335.7</td>
<td>128.0</td>
<td>205.7</td>
<td>128.5</td>
<td>10.516</td>
</tr>
</tbody>
</table>

As mentioned in the description of the tasks, the participants could decide for themselves how long they wanted to study the screen map. The last row in Table 5.1 indicates that the experts chose to study the map for a longer time than the novice users (335.7 s or 5.6 minutes vs. 205.7 s or 3.4 minutes).

5.3.2. Heatmaps

Eye movement data are often visualised using heatmaps. In Figure 5.3, four of such heatmaps from the same participant are depicted (each associated with a different stimulus). Almost all software accompanying eye tracking systems contain tools to create heatmaps. These ‘maps’ visualise the intensity levels where the participant was looking on the stimulus. Typically, a colour scale ranging over three colours is used: green (areas with lower fixation intensities) over yellow (areas with higher fixation intensities) to red (areas with very high fixation intensities). It must be noted here that in most software packages it is possible to change this colour range to a more cartographically acceptable representation – using one colour (hue) (cfr. Bertin, 1967) – but this option is rarely utilised.

Figure 5.3: Heatmaps from the same participants for the four trials

Heatmaps are an ideal tool to get an initial overview of the eye movement data, but they come with a number of serious drawback. First, it is very difficult to objectively compare the heatmaps; this is nowadays often done just at sight. Second, most
software do not allow operators to adapt the classification system. Dependent on what is under investigation, the focus of the visualisation might be on the general pattern of the fixation intensities, or on the extreme values. As a consequence, different classification schemes are desirable. However, adaptations in the standard classification scheme are often not possible. Third, it is not possible to detect extreme values between, for example, different user groups. This is again a consequence of the application of the standard classification scheme. The software determines the maximum value (total fixation duration in this case) and applies the same colour scheme on all maps based on this criterion. For example, the maximum values for the heatmaps in Figure 5.3 are respectively 6.211 s, 11.444 s, 15.445 s, and 13.048 s. Although the maximum value of the first heatmap is about half of these of the other heatmaps, the same colour scheme is applied. Consequently, from these heatmaps in Figure 5.3 it cannot be derived that the fixation intensity is much lower on the first map. Because of these drawbacks, an alternative to the heatmap visualisation is proposed in the next section: the gridded visualisation.

5.3.3. **Gridded visualisation: methodology**
A grid of AOIs was placed over each map image to obtain detailed information of the participants’ fixation in each of the grid cells. In Figure 5.4, this grid of AOIs is depicted in yellow; the cyan circles represent the participant’s fixations. Based on the size of the map image, it was decided to use square AOI’s with a size of 40 by 40 pixels. This means that a grid of 20 by 32 cells is placed over the map image, resulting in 640 AOIs. The DataViewer software (SR Research) can create so called AOI reports that list, among others, the number of fixations and total dwell time in each of the AOIs related to one trial. The AOIs report’s structure is as such that all data related to a specific AOI are represented on a single row and all AOIs are listed underneath each other. One report was created for each participant, containing the four trials.
Two columns in this report are of particular interest: the total count of fixations and the total duration of the fixations per trial, within each AOI respectively. However, as mentioned before, a significant difference in the duration of the trials was observed between the expert and novice users. As a consequence, these absolute values as such are not comparable between both user groups. The longer trials of the experts can have a significant influence on the number of fixations counted during each trial and thus on the total duration of these fixations. In order to be able to compare these measurements objectively, normalised values linked with a uniform trial duration, are used. The mean trial duration of the experts was 335.7 s, whereas this was 205.7 s for the novice users. Therefore, it was decided to use a uniform trial duration of 300 s (or 5 minutes). All data – total fixation count and total dwell time – were recalculated based on the initial and this uniform trial duration.
In order to be able to present the results visually and spatially, a program was written (in JAVA) that could read these (adapted) AOI reports and restructure the data to obtain the grid of 32 by 20 cells. All values were now placed on its correct (spatial) position of the original grid for each map. These grids were constructed for the adapted total fixation count and fixation duration, resulting in a total of 192 grids (24 participants, 4 maps, and 2 dependant variables). Next, the values in all corresponding AOIs are aggregated over all maps, separately for each user group (experts vs. novices) and each dependant variable (fixation count and fixation duration). To obtain a general overview of these data, the average value in each grid cell was calculated in the aggregated grids. Furthermore, the maximum value in each corresponding AOIs was also located to identify possible outliers and deviations in the data. These aggregation operations (average and maximum) resulted in 32 new grids (4 maps, 2 user groups, 2 variables, and 2 aggregation types).

Finally, the obtained (aggregated) values were categorised into eight different classes. A greyscale colour was assigned to each class, based on the ColorBrewer, an online tool to create usable colour scheme for maps (Brewer, 2012). The addition of this visual component facilitates the interpretation of the grids. The classification of the values was chosen as such that a general overview could be obtained regarding the location of fixation patterns and extreme values. The classification and colour scheme that was applied to the data (fixation count and fixation duration) is presented in Table 5.2.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Classification and colour schemes</th>
</tr>
</thead>
</table>
| FixCount | [0-1] | [1-2] | [2-4] | [4-6] | [6-8] | [8-10] | [10-20] | [20-...]
| FixDur   | [0.000-0.325] | [0.325-0.650] | [0.650-1.300] | [1.300-1.950] | [1.950-2.600] | [2.600-3.250] | [3.250-6.500] | [6.500-...]
| Colour   | 255 | 247 | 217 | 189 | 150 | 99 | 37 | 0 |

The fixation durations correspond to the total (summed) fixation durations for the whole trial, recalculated to the uniform trial duration of 300 s. The colour scale used to visually enhance the (spatial) presentation of these fixation duration distributions
can be keyed to the colour scale of the fixation counts. The average fixation duration for this assignment was about 0.325 s. Next, the boundaries of the fixation count classification could be recalculated based on this average. Linking the classification of the fixation durations to these of the fixation counts allows detecting regions where users are staring: the classification of the fixation duration is higher than this of the fixation count in the corresponding cell. The resulting grids are discussed in detail in the next sections.

5.3.4. **Gridded visualisation: total fixation count**
The resulting aggregated gridded visualisations are depicted in Figure 5.5 for the average values and in Figure 5.6 for the maximum values. In general, a similar pattern in the fixation counts (both for the average and the maximum values) is noticed between both user groups. This fixation pattern reflects the general structure of each of the stimuli. In the grid that corresponds with the first map, two vertical linear clusters with a higher fixation count can be identified. These correspond to the two leftmost rivers on this map. The users from both user groups focus on these linear structures, resulting in a cluster of fixations and thus a higher fixation count, to make sure that they remember a reference frame in which other map elements can be placed. The focus (or the users’ attention) on these linear elements is somewhat stronger with the experts than with the novices. Another interesting element (mostly considered by the expert group) seems to be located in the lower right corner of the map. This is a village that is flanked by a major road.

The fourth map in the experiment is the mirrored version of the first one. The structure in the fixation pattern for this fourth map could thus be the mirrored equivalent of this for the first map. In the fourth map, a similar (mirrored) pattern regarding the two leftmost rivers can indeed be identified, along with the village in the (now) upper right corner. Regarding the maximum fixation count of the fourth map (Figure 5.6), it can be noticed that the grid of the novice group has a darker background than this of the expert group. This means that the maximum values regarding the fixation count is higher in the whole map image, whereas the experts seem to direct their focus more on particular items. This pattern cannot be derived from the grid with the average values (Figure 5.5). A more detailed (statistical)
comparison of the fixation counts between map 1 and map 4 is discussed in a separate section (Section 5.3.8).

Figure 5.5: Grid of AOIs depicting the average fixation count for experts (top row) and novices (bottom row) for each stimuli (left to right)

In the grids related to map 2, a vertical linear structure can be identified in the middle of the grid. This corresponds to the location of a major road in the map. The higher fixation count related to this road indicates that this road is fixated the most during the trial, which suggests that the users especially wanted to remember this road as a reference frame. This linear structure is more obvious in the grids of the expert group, than in these of the novice group. This indicates that the experts tend to focus their attention more on this reference frame than the novices. Other focus points can be found in the top left corner and the upper right side of the map. These fixation clusters correspond to the location of the water bodies on the map, which
were visualised in a deviating colour. This deviation in colour use seems to attract the users’ attention, resulting in a higher fixation count. The deviating colour use regarding the village backgrounds does not seem to have an influence on the attentive behaviour of both user groups.

During the study, two types of stimuli were used in the third trial. Half of the participants saw the original map, the other half saw a mirrored version (over its central vertical axis). The measurements of these two stimuli are aggregated (average or maximum) and presented in one grid, related to map 3. This results in a fixation pattern that is similar (but mirrored) on the left and right side of the map. Especially the fixation count of the expert group reflects the linear structure of the vertical and horizontal road/river combination on the map. This is also present in the grids of the novices, but less pronounced. Both the experts and novices seem to have a higher fixation count on the left side of the map, although half the participants of each group saw the mirrored version. This pattern is again more pronounced in the expert group, particularly when considering the maximum fixation count. This could indicate that the map users tend to fixate more on the left side of the map.

When averaging the fixation count for each of the four quadrants (upper left, upper right, lower left, and lower right) instead of for each AOI, the fixation count is always the highest in the upper part of each map as opposed to the lower part. Furthermore, a higher number of fixations is counted in the left part of each map as opposed to the right part. More detailed information on how users looked at the map over time can be identified by studying the evolution of their scanpaths. This is discussed in Section 5.3.9.

5.3.5. Gridded visualisation: total dwell time
Besides the number of fixations a user has at a certain location, the total duration of these fixations might provide crucial insights in the users’ cognitive processes while trying to interpret the map’s content. Longer fixation durations might, on the one hand, indicate that the user finds a certain region of the visual stimulus particularly interesting. On the other hand, these longer fixation durations might also indicate that the user finds it difficult to interpret the content. This means that the user has
difficulties to recognise the object, resulting in a higher cognitive load, which is thus also linked to longer fixation durations. Two of such grids, related to map 2 for experts and novices, are depicted in Figure 5.7.

![Figure 5.7: Grid of AOIs depicting the average total dwell time (map 2)](image)

This visualisation gives a good overview regarding the fixation duration, linked with the counted number of fixations (see Figure 5.5). Similar as with the fixation counts, the general structure of the map is reflected in the grid: the main linear structures are linked with longer total dwell times, both for the experts and novices. When comparing Figure 5.5 (fixation count) and Figure 5.7, it can be concluded that the classification distribution of the fixation counts is in correspondence with this of the fixation duration. This observation holds true for both user groups and is similar for the other stimuli. In order to investigate the differences between both user groups on a more detailed level, other visualisation methods are indispensable. Two of such methods are discussed in the next sections: 3D gridded visualisation of the total dwell time (Section 5.3.6) and calculation of the average fixation duration per fixation (Section 5.3.7).

5.3.6. **3D gridded visualisation: total dwell time**

With the 3D gridded visualisations, an extra dimension is added to the original (aggregated) grids. In each cell of the grid (or AOI), a bar is constructed which height corresponds to the value in that cell. The 3D gridded visualisations of the average total dwell time are depicted in Figure 5.8 for map 1 and map 4; and in Figure 5.9 for map 2 and map 3. This visualisation allows comparing the values of the fixation durations between experts and novices in more detail, as the data is not classified. The downside of this approach is that the spatial distribution of the values on the grid is not that clear. The perspective view, in combination with the higher bars in front of the image, conceal the lower bars behind them. However, these 3D
graphs are an ideal approach to study differences in general and extreme values between two user groups, without considering their spatial location in much detail. The (normal) gridded visualisation is less suitable to obtain detailed insights in the differences of the actual values between the two user groups, but their spatial distribution is clearly visible. As a consequence, both approaches complement each other in the type of information that can be obtained.

Figure 5.8: 3D representation of the average total dwell time for map 1 and map 4

Figure 5.9: 3D representation of the average total dwell time for map 2 and map 3
Figure 5.8 presents the 3D gridded visualisation of the average total dwell time related to map 1 and map 4. The resulting graphs for both user groups are placed next to each other to allow a better comparison. Furthermore, map 1 and map 4 are placed in the same figure to be able to compare their corresponding (mirrored) values. The results of the remaining stimuli (map 2 and map 3) are depicted in Figure 5.9. These figures indicate that the resulting fixation durations in the depicted AOIs are very similar between both user groups, which is consistent with the gridded visualisations of the fixation counts. However, the novice group seems to have more extreme values – rather long fixation durations or higher bars – in each of the maps. Hardly any of the measurements related to the expert group are higher than 2.5 s; a threshold which is more often crosses by the novice group.

The extreme values related to the fourth stimulus might be explained by the fact that this map was the mirrored version of the first map. The users recognised the structure of the map, but this was also completely different (upside down). This is a typical case of proactive interference (Matlin, 2002). The users find it difficult to interpret and learn new material (map 4), because of previously learned material (map 1) that keeps interfering with the current interpretation and learning process. This negative influence on the users’ cognitive processes (a higher cognitive load) results in longer fixation durations.

The peaks observed in map 2 are, on the one hand, situated in the middle of the map, which corresponds to the location of the main vertical road on the map. The higher fixation durations on the top left and upper right side of the map, on the other hand, correspond to the locations of the two water bodies which were depicted in a deviating colour. From the fixation counts in the gridded visualisation, it could be derived that these regions were clustered with fixations, which explains these higher total dwell times. The users’ attention is attracted by these ‘strange objects’, but still higher values are observed for the novice group. However, the deviating background colour of the villages (light purple instead of light yellow) does not seem to influence the attentive behaviour of the users.

The 3D graph depicting the average fixation duration of the expert group looking at map 3 shows a more homogeneous distribution. This could be explained by the fact
that half of the participants saw the mirrored version of the map. However, the bars are higher in the left side of the map, which is in correspondence with the gridded visualisation of the fixation counts. The expert users spend more time fixating the left side of the map, despite the mirrored map image. This observation does not hold true for the novice users; the height of the bars nearly equal on the left and right side of the map. An extreme value is notice in the middle of the map, which cannot be linked to an extreme measurement in the fixation counts. This indicates that the users where staring at this location on the map. However, no special or deviation colour use or objects are located at that position on the map. The cause of this outlier can thus not be brought back to anything on the map itself.

Extreme values in the fixation duration would mean that the user fixates a certain region during an abnormal amount of time. This would indicate that the user is, on the one hand, attracted by something in that region. It can, on the other hand, also indicate regions that are difficult to interpret by the user, resulting in longer fixations but not necessary a higher count. To distinguish between these two options, an additional eye movement metric is studied in more detail: the average fixation duration of a single fixation, which can then be compared to the map’s content at that location using the (3D gridded visualisation).

5.3.7. 3D gridded visualisation: fixation duration per fixation

The average fixation duration of a single fixation was already statistically analysed in Section 5.3.3. As mentioned before, these statistical analyses miss the essential spatial element inherently linked to maps and their design. The difficulty with which a user interprets the visual content at a certain location on a map can identify problems in the map’s design. This difficulty is typically reflected in longer fixation durations, due to a higher cognitive load. However, longer fixation durations might also indicate that a certain object is more engaging in some way. The difference between both interpretations can be made by linking the (deviating) results to the actual map content at that location.

Likewise as with the total average fixation duration, a 3D gridded visualisation is created in which each bar height corresponds to the average fixation duration of a single fixation at that location. These graphs are depicted in Figure 5.10 and Figure
5.11. Comparing the results for all maps between the expert and the novice users reveals that novices tend to have more deviating (longer) fixation durations. This general trend shows that the novices find it more difficult to interpret (and thus learn) the contents of the map (for all maps), causing a higher cognitive load.

Although map 4 is the mirrored version of map 1, the difference between experts and novices is much more pronounced on map 4 (see Figure 5.10). The expert group has a number of fixation durations which are longer than average, but the novices show a cluster of very high fixation durations near the middle of the map. The position of these grid cells corresponds to the location of the calibration target that was displayed between each map to check the validity of the calibration. Consequently, this was also the region where the users were looking at the moment that the map was displayed. It can thus be derived that the novice users had longer fixation durations when the map was first displayed. This indicates confusion, which might be explained by the recognition of the first stimulus; however this is displayed ‘upside down’.

Figure 5.10: 3D representation of the average fixation duration per fixation for the expert users
The difference between expert and novice users is also clearly visible in the 3D graphs related to map 2. The higher bars in the experts’ graph are centred on the main vertical and horizontal road, whereas these of the novices’ graph are distributed over the entire map image. No particular deviations are noticed on the location of the water bodies (depicted in a deviating colour).

In map 3, the novice group is characterised by the same outlier that was also present in the 3D graphs of the total dwell times. When studying the original (not aggregated) data, it could be concluded that this high fixation duration was caused by a single participant who stared at this AOI for an extremely long time (14.2 s). This measurement distorted the average over all novice participants for this AOI. This type of ‘staring’ could be explained by the cognitive process of rehearsal in order to remember (or learn) the map image: transferring information from the WM to the LTM. The expert group also shows a number of longer fixation durations on this map, but still more are found in the novice group.

5.3.8. **Statistical grids: mirrored maps**
The statistical grids add a statistical component to the gridded visualisation. The values of all corresponding cells (AOIs) are compared statistically, using a one way ANOVA. For each grid comparison, 640 significance (P) values are thus obtained.
These results are again placed in the gridded structure. To visually present these ANOVA results, a classification scheme with four classes (and thus colours) is applied on the grid. This classification scheme is listed in Table 5.3.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Classification and colour schemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sign. (P)</td>
<td>Description</td>
</tr>
<tr>
<td>&gt; 0.1</td>
<td>not significant</td>
</tr>
<tr>
<td>[0.1-0.05]</td>
<td>near significant</td>
</tr>
<tr>
<td>[0.05-0.01]</td>
<td>significant</td>
</tr>
<tr>
<td>&lt;0.01</td>
<td>highly significant</td>
</tr>
<tr>
<td>Colour (RGB)</td>
<td></td>
</tr>
<tr>
<td>255</td>
<td>217</td>
</tr>
<tr>
<td>150</td>
<td>37</td>
</tr>
</tbody>
</table>

From Section 5.3.4 it could be concluded that the distribution of the total fixation count reflected the general structure of the map. Important and engaging items are fixated more often than other items. These important items corresponded to the main linear structures on the map, which might be used as a reference frame. What is more, the stimuli depicted during the first and fourth trial are each others’ mirrored equivalents. As a consequence, it could be expected that the patterns found in the related grids are also each others’ mirrored equivalent. The same can be expected of the fixation distributions related to the third map. Half of the participants saw the original map; the other half saw the mirrored version (this time over the vertical axis).

The method of the statistical grids is used to test whether this hypothesis holds true. The first column in Figure 5.12 depicts two of such statistical grids, related to map 1 and map 4; the second column contains the tests related to the third map. The top images depict the comparison between the original grids. The comparison between map 1 and map 4 shows a lot of significant and highly significant differences in the fixation counts of the corresponding cells in the grid. This amount is less in the comparisons related to the third map, but the map’s pattern is still clearly visible. In both statistical grids, the horizontal and vertical axis of the mirror operation can also be distinguished. This is the location where the original and mirrored version overlap, resulting in a lighter line in the statistical grids: not significantly or nearly significantly different. The $P$-values show a clear similar (mirrored) pattern on both sides of this axis (above vs. below for map 1 and map 4; left vs. right for map 3).
Figure 5.12: Statistical comparison of the number of fixation between two mirrored maps

The lower left grid in Figure 5.12 shows the statistical comparison between map 1 and the mirrored version of the grids related to map 4. By mirroring the grids related to map 4 over their horizontal axis, it could be expected that their pattern reflects the structure of the first map. This is confirmed by the depicted statistical grid. Very few (highly) significant differences are found in the grid. The majority of the grid is populated with not significant or near significant values ($P<.1$). A cluster of significant differences is found in the lower left corner of the grid, which corresponds to the location of a village and a crossroad.

An even more similar result is obtained when comparing the grids of (the original version of) map 3 with the mirrored grids of the adapted third stimulus (lower right corner in Figure 5.12). In this grid, almost no highly significant differences are found. The better result regarding the third stimulus might be explained by the fact that the participants saw both the first and fourth map, which caused confusion (particularly among the novice users). The two version of map 3 were looked upon by two separate user groups, avoiding influences on the cognitive processes due to recognition or confusion. The statistical grids at the bottom of Figure 5.12 thus indicate that the patterns of the users’ fixations are guided mainly by the main (linear) structures on the map.
5.3.9. **Scanpath visualisation**

Another way to visually explore and analyse the spatial dimension of the eye movements is to study the scanpaths of the participants. These scanpaths are sequences of subsequent fixations and saccades. The Visual Analytics Toolkit was used to visualise the participants’ scanpaths on top of the actual stimuli. Filter operations based on attributes (stimuli and user group) and on time intervals facilitate the visual analyses of the eye movements (see Chapter 3). Figure 5.13 illustrates these scanpaths, separately for each map and user group. What is more, to study the evolution of these scanpaths, different time intervals (during the first minute) are also depicted: 0 to 10 s; 0 to 30 s; 30 to 60 s.

The location of the participants’ scanpaths seems to be clustered on the main structuring elements of the map (major roads and rivers), with a very similar pattern between the experts and novices. This pattern remains visible during the entire first minute: the participants keep directing their attention on these main (often linear) elements. During the second half of the first minute, the participants also fixate other objects, but the structuring elements still receive much attention.

When comparing the scanpaths during the first ten seconds between the first and fourth map, a striking difference is observed. In the first map (similarly as with the other maps) the participants already direct their attention on the main structuring elements in the map. This holds especially true for the participants in the expert group. However, the scanpaths associated with the fourth map are rather chaotic during the first ten seconds. The structure of the two leftmost vertical lines is not present, which was the case for the first map. Especially the novice map users show a high number of horizontal scanpath lines, zigzagging across the image during the first 30 seconds of the map’s display. These chaotic scanpaths indicate that the users are confused by this mirrored map image.

The experts’ scanpaths on map 2 show a cluster on the horizontal and vertical major road in the map image during the first ten seconds. The novices seem to be more distracted by the water bodies (with the deviating colour use). In the first 30 s, the expert users also fixate more on these water bodies, but again less during the second half of the first minute. The villages at the bottom (left) receive also more attention.
during this latter interval. The deviating colour use to depict the background of the villages does not seem to influence the participants’ attentive behaviour.

The scanpaths on map 3a and map 3b are very dislike during the first ten seconds, although the same structures are found in the map image, which is especially striking for the expert users. On map 3a the experts focus on the vertical river/road on the left side of the image and less on the horizontal road/river. On map 3b the focus is more situated on the horizontal main linear structure and not so much on the vertical road/river on the right side of the map. This pattern is still visible in the longer time intervals and corresponds with the findings of the fixation counts and fixation durations. The users, particularly the experts, tend to be more attracted to the left side of the map.
Figure 5.13: Evolution of the participants’ scanpaths for the different maps and user groups.
5.4. Discussion and conclusion

The study described in this chapter is an extension of the work described in Chapter 2, in order to verify whether their results could be generalised to wider array of, and thus more complex, map types. The main aim of the experiment is to gain a better understanding in how expert and novice map users process and interpret the (complex) visual information on maps. Furthermore, deviations in the stimuli are introduced in order to study their effects on the map users’ eye movements, and thus on their attentive behaviour.

The statistical analyses confirm the findings of Chapter 2. The experts’ fixation durations are significantly shorter than these of the novices. This indicates that their interpretation process, including the different stages of object recognition, is much faster than this of the novice users. This could be explained by the level of experience, and thus amount of background knowledge stored in the LTM, that the expert users have in comparison to the novices. Due to these shorter fixations, the experts can have more fixations per seconds, and can thus interpret a larger part of the map in the same amount of time. As a consequence, it can be concluded that expert map users can interpret the maps more efficiently, both when a simple and complex maps are considered.

In order to spatially analyse these results, different approaches are presented, which all complement each other: gridded visualisations, 3D gridded visualisations, statistical grids, and scanpaths analyses. These visual and (mainly) qualitative methods shed light on different aspects of the user’s interpretation process and allow studying differences in the attentive behaviour of expert and novice users. What is more, a number of eye movement metrics related to the users’ fixations are analysed and compared: total fixation count in one trial, total dwell time in one trial, and average fixation duration of a single fixation. These analyses take the spatial distribution of the fixations across the map image into account, and are based on a grid of square AOIs.

From these analyses it could be concluded that both user groups focus their attention on a reference frame in the map image, resulting in a higher number of fixations and thus longer total dwell times. This reference frame mostly consists out
of major linear structures, such as roads and rivers. As a consequence, the main structure of the map is reflected in the gridded visualisation of the total fixation counts and durations. The focus on these linear structures is more pronounced in the grids of the expert group than with the novices. The visualisation of the users’ scanpaths during the first ten and thirty seconds also shows that the user’s attention is immediately directed towards these structuring elements. Nevertheless, the novices’ eye movement measurements show more extreme values than the experts regarding the number and duration of the fixations.

These findings correspond to the Gestalt Laws regarding the tendency of humans to organise or group the visual information on a scene (such as maps) (Wertheimer, 1923). Kulhavy and Stock (1996) stated that the visual images of maps are organised in feature and structural information. They define the structural information as the spatial framework within which the other map objects (features) can be placed. Furthermore, Huynh and Doherty (2007) showed that, when drawing a map from memory, the paths were drawn more frequently at first. This indicates their importance in the users’ memory structures and thus for the interpretation process.

Wolfe (1994) described and evaluated a modified model for guided search that contains rules which guide humans’ attentive behaviour during a visual search and Hegarty, et al. (2010) found that map users’ attentive behaviour can be modified by the salience of map objects. The results from the experiment described in this chapter indicate that the deviating colour use of the water bodies in the second stimulus also has an influence on the users’ attentive behaviour. Both user groups are attracted by these objects, which is reflected in the location of the eye movements. A higher number of fixations are found in the top left and upper right region on the second map. However, this attraction seems to be stronger in the novice group than with the experts. This latter user group focuses more on the central horizontal and vertical reference frame (major roads) in the map image.

Organising and interpreting a scene is based on previous experiences. The current visual input is linked with knowledge stored in memory through top-down processing. This also closely linked with the Gestalt Law of experience (MacEachren, 1995). However, in the case of the water bodies, the shape might look
familiar, but the colour is not. This is a discrepancy that the user has to solve, which disturbs the interpretation process. This disturbance in the interpretation process can be observed as confusion. Nevertheless, the deviating colour use of the villages’ background does not seem to have any influence on the users’ attentive behaviour or interpretation process. These objects are not part of the (group of) structuring elements on the map and, consequently, contribute less to the interpretation process of the whole image.

Two types of maps were used for the third stimulus: half of the participants saw the original map; the other half saw the mirrored version (over its central vertical axis). The superimposed results show a mirrored pattern in the total fixation count and fixation duration. However, more fixations (and thus higher total dwell times) are found on the left side of the superimposed result. This indicates that the users tend to fixated more on the left side of the map, regardless the content of the map. A quadrant analyses on all maps also indicates that this also holds true for all other stimuli presented during the study. This observation is more distinct in the measurements of the expert group. Nevertheless, based on the findings of the registered eye movements it could also be concluded that the users’ POR are mainly guided by the main structures on the map. When mirroring the results of the adapted map 3 again over its central vertical axis, the fixation clusters corresponded to these of the original map and thus to the general structure of the original map.

These results provide evidence that the subsequent foci of attention are mainly guided by the content of the image: a mirrored content results in a mirrored pattern in the scanpaths. Furthermore, the main structuring elements are clearly reflected in these scanpaths. As a consequence, it should be possible to define, and thus predict, these strategies in a model, as was proposed by Wolfe (1994).

This latter observation also holds true for the results related to map 1 and map 4. The mirrored version of the results related to map 4 corresponds to these of map 1. Nevertheless, the statistical grid showed more significant differences in the corresponding cells than with the third stimulus. This could be explained by the fact that the participants saw both the original map and the mirrored map during the study. Other results related to map 4 also show evidence of proactive interference.
The map users are confused by this map because it is very similar to information they have processed before. However, the visual information is also significantly different because it is upside down. These confusions and difficulties in the interpretation process translate into longer fixation duration, which is especially visible in the novices’ their eye movement recordings.

In the gridded visualisation of the total fixation count related to map 4, no obvious deviation values are detected. However, the 3D gridded visualisation of the total dwell time shows a number of extreme values. These values are explained in the 3D graphs representing the average fixation durations for a single fixation, where also a number of peaks are observed. This confusion can also be derived from the scanpath visualisations. During the first 30 seconds, eye movements are not immediately directed towards the structuring elements (as was the case for the first map). The scanpaths show a more chaotic distribution over the map image.

These findings indicate that the users encounter difficulties in the interpretation process due to knowledge previously stored in memory. This memory is addressed because of its familiarity or correspondence with the current visual stimuli. However, it is also essentially different, which complicates the interpretation processes. As a consequence, it can be concluded that the top-down influence is not always beneficial in the interpretation process.

The results described in this chapter give insights in how participants (experts and novice) organise and group the visual content of maps which guide their attentive behaviour and thus the interpretation process. Previous research has indicated that a good map design can facilitate the performance of this interpretation process (Hegarty, et al., 2010). The focus of the users’ attention can be influenced by the design of the objects (such as their salience). The results described above indicate that users’ tend to focus on main structuring elements in the map image. In order to facilitate the interpretation process, the objects that contribute to this main structure should thus be clearly visible at first glance, for example by increasing their saliency. Furthermore, users’ are influenced by deviations from what they are familiar with regarding these structuring elements. It can thus be recommended that the design of these objects should not be altered significantly as this may cause
confusion and thus problems in the interpretation process. Finally, the map users’ interpretation process is not influenced by deviations in the design of elements that do not contribute to the main structure of the map layout. The consequence of this is twofold. First, their design can be altered more easily, without disturbing the map users’ cognitive processes. Second, their design does not have to be optimal as its influence on the interpretation process is minimal.

References


Chapter 6

Listen to the Map User: Cognition, Memory, and Expertise

Modified from:
Ooms, K., De Maeyer, P., & Fack, V., Understanding expert and novice map users: Cognition and memory. (Submitted to The Cartographic Journal)

This chapter aims at extending current research regarding map users’ cognitive processes while working with screen maps. Special attention is paid to how information is retrieved from memory, focusing on experts and novices. A user study was conducted in which participants had to draw a map from memory. During this task, they were instructed to say out loud every thought that came into mind. Both user groups addressed the same general cognitive structures and processes to solve the task at hand. However, the experts’ background knowledge facilitated the retrieval process and allowed them to derive extra information through deductive reasoning. The novice users used more descriptive instead of naming the objects and could remember less, and less detailed map elements.

Keywords: cartography; user study; thinking aloud; cognition; memory
6.1. Introduction
As a consequence of the accessibility of the Internet, more and more novices started to use cartographic products on the web. However, because of the limited background knowledge and expertise in interpreting and working with these – often dynamic and interactive – screen maps, it might be difficult for these novice map users to process their content. What is more, maps found on the web nowadays are not only created by trained cartographers. With the increasing user involvements of Web 2.0, novice map users can contribute to the content of web maps (e.g. crowdsourcing and VGI) and can even – fairly easily – create cartographic products themselves (e.g. mashups, push pin maps, APIs). Since these ‘new map makers’ often did not receive any cartographic training, the created products might have issues regarding their design, resulting in a less effective communication process (Cartwright, 2012; Haklay, et al., 2008; Turner, 2006).

At the beginning of the new century, a number of authors expressed their concerns regarding the cognitive limits of map users in the light of modern geovisualisation techniques (among others, MacEachren & Kraak, 2001; Slocum, et al., 2001). An enormous amount of cartographic products have been released on the Internet without prior evaluation by representative users or without even considering these users’ needs. As a consequence, almost no information regarding the influence of these screen based visualisations on the users’ cognitive processes is available. A number of authors have already answered this need, especially in the case of (geographic) mobile applications (Haklay & Nivala, 2010) and navigation systems (Muenzer & Stahl, 2011; Schlender, et al., 2000).

Recently, Montello (2009) and Fabrikant and Lobben (2009) pointed out that, although a rise in the research regarding the cognitive issues (and limits) of map users has been noticed, it is still a research challenge that needs attention. Detailed insights in the users’ cognitive processes, while working with digital screen maps, are essential to be able to create effective digital maps (with their interactions, dynamic responses, and animations) in the future. However, one important issue that is often neglected in this is the influence of (cartographic) expertise on the users’ cognitive processes. The experiments described in this article focus hereby on the processes addressed while retrieving information from memory, previously gathered from a
topographic screen map. An overview of the state of the art regarding research on human’s cognitive and memory structures is given in the next sections.

6.1.1. Understanding the map users’ cognitive structures
In order to understand how users store and retrieve information, a number of memory structures need to be explained: the sensory stores, the working memory, and the long term memory. The sensory stores receive the information from our senses, including vision. A part of this information is transferred to the working memory (WM), which was also called short term memory in the past. This working memory contains the information that is actively used, but has a limited capacity. Nevertheless, there is still some debate whether there is also a time limit on the information contained in the WM (Cowan, 2001; Ericsson & Simon, 1980; Miller, 1956). In order to be able to remember the information contained in the WM for a longer time span, it has to be transferred to the long term memory (LTM). The transmission of information from the WM to the LTM corresponds to the learning process and is accomplished through rehearsal. The LTM has a virtually unlimited capacity, but cannot be reported from directly. In order to actively use the information stored in the LTM, it has to be retrieved again: transmission from the LTM to the WM (Cowan, 2001; Matlin, 2002).

General knowledge regarding a certain item, event, or situation is coded into the LTM using schemas. These schemas describe the essential concepts or characteristics that define such an item, event, or situation in a hierarchical way. Furthermore, concepts in these schemas are interconnected (through links or pointers) with related concepts in other schemas. These schemas are used during the cognitive process of object recognition and assist in the selection of certain memories stored in the LTM (MacEachren, 1995; Matlin, 2002).

When these schemas contain knowledge regarding the spatial layout of certain features, this is often referred to as a cognitive map or mental map (Downs & Stea, 1997). Cognitive maps can be constructed from two different sources of knowledge. Primary knowledge is gathered when humans view or move through the environment itself. This is not the same information that is gathered when studying a map, which is a secondary source of information and has picture-like properties.
Kulhavy and Stock (1996) argued that these visual images of maps are described by features on the one hand – characteristics of point locations, needed to identify them and their visual variables as proposed by Bertin (1967) – and by structural information – the spatial framework and relations between the objects – on the other hand. They indicate that the structural information is a critical element of cognitive maps. Other authors consider the cognitive map to consist of nodes (or landmarks) and paths, which form a hierarchical structure in the associated schemas (e.g. Hirtle & Jonides, 1985; Huynh & Doherty, 2007). However, most research on the structure of human’s cognitive maps was conducted in relation to urban environments when primary knowledge was gathered.

6.1.2. Influence of expertise

Miller (1956) stated that the WM can hold only seven ‘chunks’ of information, give or take two. More recent research found that this number is even less and corresponds to four chunks of information, give or take one (Cowan, 2001). These chunks can be defined as groups of information that have a strong internal relation, but weaker relations to other information. Chunking of information is closely linked to the structure of the schemas held in the LTM. Nevertheless, this limited number of chunks that can be contained in the WM does not say anything about the actual amount of information. Related information can be grouped together, forming only one larger chunk of information (corresponding to one ‘slot’ in the WM). Effective grouping of information can thus increase the amount of information stored in the WM (Gobet & Simon, 1998; Miller, 1956).

Gobet and Simon (1996, 1998) studied the cognitive processes and memory structures of expert chess players. They concluded that the better performance of these experts can, on the one hand, be explained by the higher amount of chunks that are stored in the LTM. These chunks of information are, on the other hand, also much larger: more information can be contained in a single chunk (Matlin, 2002). Because the experts have richer schemas stored in memory, they can structure this information better and more tightly in the chunks that are transferred to the WM. Similar results were found in other domains of expertise, such as bridging and computer programming (Gilhooly, et al., 1998).
Based on theories of human cognition, Thorndyke and Stasz (1980) assumed that the descriptions stored in memory are based on existing knowledge and the available cognitive processes. This is closely linked with the findings of Hambrick and Engle (2002) that domain knowledge has a significant (positive) influence on a person’s memory performance. They suggested that domain knowledge can serve as a retrieval structure that facilitates the retrieval process from memory. Kulhavy and Stock (1996) further distinguish between general map knowledge and specific map knowledge. They argue that experts and novices differ little in the amount of information that can be retrieved from memory when standard map information is considered. The specific map knowledge differs more between experts and novices and, consequently, influences the cognitive maps that are constructed when interpreting maps.

In their experiments Thorndyke and Stasz (1980) and Gilhooly, et al. (1988) investigated the difference between expert and novice map users in how they learn and recall the information presented on planimetric and contour maps. Besides the cognitive processes, they focussed on the different strategies that were addressed during learning and information retrieval. They found that experts did not perform better in recalling the planimetric maps and the non-contour features on the contour maps. However, they could recall (and draw) the contour information from the contour maps better. The authors concluded that experts retrieved both specialised and ‘lay’ or lower level schemas. Furthermore, the experts’ lower level schemas are richer than these of the novices (although the same labels were used) and they could link these lower level schemas to the specialist schemas. These findings were also confirmed by Montello, et al. (1994), who used landscape scenes and topographic maps in their experiments.

In Chapter 2 the results indicated that expert map users are significantly more efficient in the performance of a visual search on a basic screen map design. The experts’ higher efficiency in map interpretation was further confirmed in Chapter 5 when both expert and novices map users had to learn – store in memory – the content of topographic maps. Nevertheless, the communication of spatial information through maps is only successful if this information can be recalled and used later on. Still very little is known on the cognitive processes and strategies that
are used to recall the information that was retrieved from these screen maps. The aim of this paper is to study the cognitive processes and strategies involved in remembering (recalling) the information that was gathered previously from screen maps. The techniques used to study the cognitive processes during information retrieval are described in the next sections.

6.1.3. Accessing the users’ cognitive processes
6.1.3.1. Thinking aloud
Nielsen (1993, p.195) stated that thinking aloud could be “the single most valuable usability engineering method”. Thinking aloud is also a well integrated method in psychological research. During a thinking aloud study the participants are typically asked to say out loud every thought that comes into mind. In order to maintain objectivity, the participants should not interpret these thoughts, only verbalise them. As a consequence, these verbal protocols are direct and unfiltered as opposed to other methods such as retrospection and introspection (Nielsen, 1993; Trickett & Trafton, 2007; van Someren, et al., 1994). Since humans can only report information contained in the WM (See Section 6.1.1), the verbal protocols obtained during concurrent verbalisation correspond to the content of the WM at that moment (Trickett & Trafton, 2007; van Someren, et al., 1994).

In the past, some concerns regarding the validity and reactivity of the method have been expressed. The validity of the method, on the one hand, is related to how well the obtained verbal protocols reflect the user’s thoughts (and cognitive processes). Ericsson (2006), among others, concluded that the validity of the verbal protocols is dependent on the time interval between the occurrence of the actual thought and its verbalisation. Consequently, concurrent verbalisations show the highest level of validity (and completeness). What is more, a considerable amount of the participants’ cognitive load is already addressed to solve the assignment and consequently, there is no room left to interpret or change the thoughts before or during verbalisation (van Someren, et al., 1994).

The reactivity of the method is, on the other hand, related to influences that the verbalisation may have on the user’s thoughts (the sequence of the thoughts, for example) and the execution of the assignment. Ericsson and Simon (1980) found no
influence due to the application of the thinking aloud method on the users’ thoughts, in comparison to participants who did not verbalise their thoughts. However, the participants might need more time to complete the assignment when they have to verbalise their thoughts in addition to the assignment (Ericsson, 2006). As a consequence, thinking aloud studies should not be combined with time measurements.

An excellent overview of the methodology regarding the analyses of verbal protocols is provided by van Someren, et al. (1994) and Chi (1997). The structure of these analyses is depicted in Figure 6.1. Based on the task analysis and the existing psychological theories (such as described in Section 6.1.1 and Section 6.1.2), a psychological model can be constructed. This model depicts the different cognitive processes that are needed to complete the task, and the relation between these different processes. A set of standard codes can be linked to each of these cognitive processes, possibly on different levels of detail. The obtained (raw) verbal protocols (transcriptions) need to be segmented or divided into smaller units. A code can be assigned to each of these units, indicating which cognitive process takes place at that moment. The results of this latter process are the coded protocols, which list objective information regarding the participants’ cognitive processes while executing the assignment (Chi, 1997; van Elzakker, 2004; van Someren, et al., 1994).

Suchan and Brewer (2000) consider qualitative methods in combination with content analysis suitable for research on map making and map use. One of the methods they propose is the collection of verbal data through thinking aloud. The experiments by, for example, Thorndyke and Stasz (1980), Gilhooly, et al. (1988), and Montello, et al. (1994) (mentioned in Section 6.1.2) used, among others, thinking aloud to study differences in how expert and novice map users learn and recall information from paper maps (planimetric, contour, and topographic maps).
6.1.3.2. Sketch maps

Huynh and Doherty (2007, p.286) defined sketch maps as “the extraction of information from a mental map through drawing”. On the most abstract level, Blaser (2000) defined sketches as a collection of line strokes that can be grouped together to form objects and these are in turn related to other objects. He distinguishes between three main parts that build up a sketch map: sketched objects, spatial relations, and annotations. In his article, Blaser (2000) described the results of a survey that investigates which objects, spatial relations, and annotations are typically drawn while sketching a map.

Huynh, et al. (2008) and Huynh and Doherty (2007) focussed on the order with which the sketched objects were drawn. They distinguish between two elements that can be drawn - paths and landmarks – and conclude that paths are more frequently drawn at first; landmarks are more frequently drawn later on. However, they focussed on cognitive maps from urban environments (primary knowledge), not on information gathered from maps (secondary knowledge). Thorndyke and Stasz
(1980), Gilhooly, *et al.* (1988), and Montello, *et al.* (1994) combined thinking aloud with sketch maps. They found that experts only performed better in the sketch maps when specific map knowledge had to be addressed: contours and topographic maps.

### 6.1.4. Research objective and question

The aim of the experiment described in this article is to investigate how map users retrieve information from memory that was previously gathered from screen maps. In order to use and apply this information later on, the users have to be able to retrieve this again from memory. Since both experts and novices in cartography nowadays make and use the screen maps that can be found on the web, it is important to study the difference between these target groups. Experts and novices have to be able to create (screen) maps that take into account the cognitive limits of their target users.

In Chapter 2 and Chapter 5 the eye movements of (expert and novice) map users who interpreted – stored in memory – the content of screen maps with a simple and a complex content were examined. This chapter aims at extending this research by investigating the next step in the communication process: the retrieval of information from memory previously gathered from screen maps. The main research question related to this objective is to investigate whether experts and novices address different cognitive processes and/or use different strategies to recall the information. The set-up of the experiments is based on these conducted previously by Thorndyke and Stasz (1980), Gilhooly, *et al.* (1988), and Montello, *et al.* (1994) who studied these processes and strategies using printed planimetric and contour maps. The combination of two methodologies will be used during the experiment, in order to provide complementary data. Thinking aloud will be used to study the users’ cognitive processes and strategies during information retrieval. However, the spatial and visual nature of the information that is recalled can be difficult to verbalise (Kulhavy & Stock, 1996). Therefore, the resulting sketch maps are also analysed regarding their content.
6.2. Materials and methods

6.2.1. Participants
Since the analyses of thinking aloud data are very labour intensive, the number of participants is preferably limited. However, a sufficient amount of users from each group (experts and novices in this case) are also indispensible to be able to detect differences between both. Nielsen (1994) recommended to select four (plus or minus one) participants. He based this finding on the ratio of usability problems that could be discovered with a certain number of participants. Tricket and Trafton (2007) concluded that it would be ideal to have more than five participants (preferably five to ten).

A total of 24 participants took part in the study described in this chapter, with 12 experts and 12 novices. Both groups consisted out of six female and six male participants. It was chosen to work with 12 participants for each group (more than necessary) because some stimuli were only shown to half of them. The participants that took part in the expert group had at least a MSc. in Geography or Geomatics. Furthermore, they were employed at the Department of Geography (Ghent University) at the moment of the study and work with cartographic products on a daily basis. The novice map users did not receive any previous cartographic training and do not work with cartographic products on a daily basis. All participants’ native language was Dutch, the language in which the study was conducted (explanation of the assignment, thinking aloud, etc.).

6.2.2. Tasks and stimuli
Four different maps were displayed on a screen to the participants during the experiment. They were instructed to remember the map that was currently displayed as best as possible in order to be able to draw it afterwards. They were told that they did not have to remember the location of each individual house, but certainly the main structuring elements on the map. The participants could decide for themselves how long they wanted to study each screen map, with a maximum of ten minutes. This latter time limit was introduced to keep the length of the study manageable, but this was rarely reached. Most participants ended this initial part of the trial after (on average) five minutes. During this initial part, the participants’ eye movements were
recorded. A detailed description of these eye movement recordings can be found in the previous chapter (see Chapter 5).

The stimuli that were used during the experiment are depicted in Figure 6.2. These maps are selected out of the Belgian topographic map series on 1:10 000, based on the criteria that they were not overly complex, contained a number of main structuring elements, and did not contain any well known areas or cities. Some extra deviations were introduced in these map images. Half of the participants saw map 3a, whereas the other half saw map 3b (with an equal share of experts and novices). These two maps are each others’ mirrored equivalents (over its central vertical axis). The map displayed during the fourth trial was the mirrored version of the one displayed in the first trial, this time over its central horizontal axis. Finally, the colours of some objects in the second map – the water bodies and the village backgrounds – were adapted to obtain a non conform result.

After ending the display of the screen map, the participants were asked to take place at another table where they could draw the map which they had just studied. They were instructed to say out loud every thought that came into mind, even if it did not have to do anything with the experiment itself. No time limit was set on this drawing assignment to avoid any pressure on the participants that could influence their behaviour. The absence of a time limit is also important as the participants are asked to execute two concurrent tasks: drawing and verbalising their thoughts. These verbalisations influence the drawing times, and these measurements are – as a consequence – not used in the analyses.

On the table, participants could find a pile of blank A3 pages on which they could draw the map. Twelve colour pencils, a sharpener, and an eraser were placed at the upper side of the A3 pages. The colour pencils were arranged in the same order before the start of each experiment, but were not touched by the experiment moderator between the trials. On the right side of the A3 pages, a coloured post-it was placed on which the participant’s code was written (e.g. ef_p3: expert, female, third participant). The colour of the post-it corresponded to the trial number (trial 1 = green; trial 2 = yellow; trial 3 = orange; trial 4 = pink). This trial number could thus easily be derived from the recorded videos (see Section 6.2.3). The participants
were informed that they could start drawing (and verbalising their thoughts) once
this post-it was in place.

Figure 6.2: The stimuli presented on screen during the initial part of the trial

In total, the participants had to complete four trials of remembering and drawing a
map. The assignment was read out loud to each participant, and the same text was
also displayed on the screen before the start of the first trial. At this point the
participant could ask any question if the assignment was not clear. These
instructions also informed the participants that they had to fill in a post study
questionnaire after the completion of the fourth trial. This questionnaire was
necessary to check whether any of the participants were familiar with the displayed
regions, to register personal characteristics, and to receive feedback.
6.2.3. **Apparatus and recordings**

The experiment was conducted in the Eye Tracking Labo of the Department of Experimental Psychology (Ghent University). This laboratory is standard equipped with an *EyeLink1000* (SR Research) to record participants’ eye movements while studying stimuli on a screen. The stimuli (maps in this case) were depicted on a 21 inch monitor. Furthermore, two camera’s and a headset were used to record the essential auditory and visual components related to a thinking aloud study. The participants had to put on a headset (with a microphone) before the start of the second phase of each trial: drawing the map. This headset was connected to a laptop which webcam recorded the front view of each participant: face and A3 paper on which they were drawing. This was an *Acer Crystal Eye Webcam* that captured HD videos, including the audio recorded with the headset.

A second HD camera (*Sony HandyCam HDR-CS115E*) was placed on a tripod and captured the drawing process on the A3 page from a top view. The coloured post-it placed next to each A3 page was visible in all recorded videos. Furthermore, the code assigned to each participant was written on the post-its and could thus be read from the top view videos (from the HandyCam). The synchronisation of the recordings from the two cameras (webcam and HandyCam) was based on the moment in time when the post-it was placed (more specifically, the moment when the finger of the moderator did not touch the post-it anymore). The software ELAN (*EUDICO Linguistic Annotator*, developed by the Technical Group of the Max Planck Institute for Psycholinguistics) was used to manually synchronise the videos and prepare them to be able to analyse the derived verbal protocols. These analyses are described in more detail in Section 6.2.4.

6.2.4. **Methodology**

The recorded videos were analysed based on the methodology described by, among others, Chi (1997), van Elzakker (2004), and van Someren, *et al.* (1994). This methodology was also briefly described in Section 6.1.3 and illustrated in Figure 6.1. A number of subsequent steps are necessary to obtain the final coded protocols in an objective way. First, the obtained videos have to be transcribed, which will form the raw (verbal) protocols. Second, the obtained verbal utterances have to be divided into smaller units (to which the codes will be assigned). This is described in
Section 6.2.4.1. Next, a suitable coding scheme has to be developed, based on a psychological model (see Section 6.2.4.2 and Section 6.2.4.3). Finally, these codes will be assigned to all segmented protocols (see Section 6.2.4.4).

6.2.4.1. **Segmentation of transcriptions**
Chi (1997) described a number of issues that have to be considered when choosing the ‘size’ of these units. She concluded that there is a trade-off between the segmentation granularity and the information that can be retrieved. Smaller sized units also have a severe impact on the amount of work needed to analyse them. Combining two different segmentation types would result in redundant information (which can be used as a validity check), but could also result in complementary information. Smaller units might capture inferences that could not be derived from the coarser segmentations. The size of these units should also be linked to the actual research question. Furthermore, the identification of the different segments can be language based (sentences) or activity based (pauses). Finally, sometimes the occurrence of only certain (combinations of) words or activities is of interest. In this case it is not necessary to segment the whole transcription, but searching for these words or activities would suffice.

Related to the main research question of this article (how do different types of map users mentally store and retrieve the visual information presented to them) two segmentation sizes are selected: sentences and words. These two levels are chosen to be able to check the validity of the data, on the one hand, and to be able to obtain complementary (more detailed) information on the other hand. Segmentation on word level allows obtaining very detailed information regarding the actual use of wording: the description of their thoughts. These word segmentations might give insights in how the information is actually stored: the pointers that were used to trigger (activate) information in the LTM, actual description of the visual information stored in the WM, etc. The sentence segmentation corresponds to the structure of the users’ full thought. Using this granularity of segmentation allows maintaining the contexts in which certain words are used. Furthermore, the size of a ‘full thought’ allows assigning codes to each unit that corresponds to a certain cognitive process, which is less obvious on the word level.
The full thoughts can be marked both by linguistic elements (sentences) and actions (longer pauses). For the word segmentation, each individual word that was spoken during the study was listed. In total 2,043 different words were identified. This initial word segmentation was aggregated by grouping words that have the same meaning, such as for example forest and woodland or buildings and houses. This aggregation operation resulted in a list of 1,189 words. Next to each word in the aggregated result, its total number of occurrences was listed, separately for each user group.

6.2.4.2. Task analysis and psychological model

The psychological model is based on the task that the participants had to complete. Based on current psychological theories (See Section 6.1.1 and Section 6.1.2), the cognitive processes that would be addressed during the execution of this task are identified, including their interrelations. The participants’ assignment was simple and clear: draw the map that you have just seen on an A3 page using colour pencils. On the highest (general) level, independent of the actual assignment, this task could be modelled as depicted in Figure 6.3. The participant first orientates himself and his surroundings (what material is available, how do I begin, etc.). Next, he starts to execute the assignment and at a certain moment he evaluates his temporary result. If the participant finds that this result is acceptable, the assignment is completed. Otherwise, the participant will correct the result by adding or adapting elements, until he will evaluate this (new) result again. The ‘Execute Assignment’ and ‘Evaluate Result’ sub-parts of this model will be described in more detail in the next paragraphs.

![Diagram of the psychological model](image)

Figure 6.3: Highest (general) level of the psychological model

Figure 6.4 depicts the psychological model related to the actual assignment execution. At this level, links with the participants’ cognitive processes are made, including the relations between these processes. First, the participant will consult his
WM whether this contains information (map elements) that still needs to be drawn. If so, these elements will be drawn. If not, the participant will have to consult his LTM (through pointers to schemas) in order to retrieve this information: LTM to WM. Information has to be transferred to the WM to be able to draw it. Next, the item that was just drawn will be evaluated for correctness and possibly adapted. If the participant finds the drawn item acceptable, he will consult his WM again. This process will continue until the participant evaluates the total image that was drawn hitherto. This evaluation process is modelled in Figure 6.5.

Figure 6.4: Psychological model of the assignment execution

During the evaluation process, the participant will look at the image he had drawn hitherto. This image will be ‘captured’ through his (visual) sensory store and transferred to the WM. The memories of the original screen map are stored in the LTM and have to be retrieved again. Next, the drawn image is compared with the memories of the original image, based on the shapes, location, completeness, colours, distributions, etc.

Figure 6.5: Psychological model of the evaluation process

Figure 6.6 models the information retrieval process, from LTM to WM, in more detail. In order to be able to retrieve chunks of information from the LTM, this information has to be ‘activated’ to create the so called Long Term Working
Memory (LTWM) (Cowan, 2001). This activation and retrieval process makes use of pointers to the schemas (or links between these schemas), stored in the LTM. These pointers work as triggers to activate and subsequently retrieve the correct (chunks of) information from the LTM. If the correct, usable, information is retrieved, the chunk of information is placed in a slot of the WM. As mentioned before, the size of these chunks can vary containing, for example, the name of an object, its shape, location, related background information, etc.

![Figure 6.6: Psychological model of the information retrieval process (from LTM to WM)](image)

The combination of the previous schemes results in the total psychological model, related to the experiments’ assignment. This model is based on the psychological theories that are currently generally accepted and which were previously described in Section 6.1.1. Based on this model, a suitable coding scheme can be developed. This coding scheme is described in detail in the next section (Section 6.2.4.3).

### 6.2.4.3 Coding scheme

The coding scheme proposed below is based on, and inherently linked with, the psychological model described in the previous section (Section 6.2.4.2). A code will be assigned to each process that is addressed during the completion of the assignment (and thus described in the model). In order to keep these codes surveyable and manageable, different levels are introduced. The codes addressed in Level 1 to Level 4 are linked to the sentence segmentations. Consequently, each segmented protocol will receive four different codes (one of each level). The codes of Level 5 are used to analyse the word segmentations.
In Table 6.1 to Table 6.5, the proposed codes, their meaning, their description, and link to the psychological model are listed. The codes are an abbreviation of the proposed categories, with a suffix (number) indicating the level of the code. The codes of the first level, the map level, correspond to the psychological model depicted in Figure 6.3 and are listed in Table 6.1. The codes depicted in Table 6.2 are linked to the cognitive processes which are modeled and depicted in Figure 6.4 (Execute Assignment).

### Table 6.1: Proposed codes related to the map level (Level 1)

<table>
<thead>
<tr>
<th>psych. model</th>
<th>category</th>
<th>code</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orientate</td>
<td>orientate</td>
<td>or1</td>
<td>before the start of the task execution, gather thoughts, check tools, assignment, ...</td>
</tr>
<tr>
<td>Execute Assignment</td>
<td>Execute</td>
<td>ex1</td>
<td>drawing, talking, explaining, correcting items, ...</td>
</tr>
<tr>
<td>Evaluate Result</td>
<td>Evaluate</td>
<td>ev1</td>
<td>check if everything is on the map (count objects, check size, relative positions,...); not the correction of individual items when drawing them</td>
</tr>
<tr>
<td>OK?</td>
<td>Evaluate</td>
<td>ev1</td>
<td>OK?</td>
</tr>
</tbody>
</table>

### Table 6.2: Proposed codes related to the item level (Level 2)

<table>
<thead>
<tr>
<th>psych. model</th>
<th>category</th>
<th>code</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consult WM</td>
<td>gather thoughts</td>
<td>gt2</td>
<td>participant is clearly thinking, gathering thoughts, finding links between items, consult previous knowledge, similarity between objects</td>
</tr>
<tr>
<td>Next Item</td>
<td>gather thoughts</td>
<td>gt2</td>
<td>items are draw on paper, colour pencil was taken, ...</td>
</tr>
<tr>
<td>LTM to WM</td>
<td>gather thoughts</td>
<td>gt2</td>
<td>check correctness of item that was drawn, or items related to it</td>
</tr>
<tr>
<td>Draw Item</td>
<td>draw</td>
<td>dw2</td>
<td>correct items which were evaluated (drawing, erasing)</td>
</tr>
<tr>
<td>Evaluate Item</td>
<td>evaluate</td>
<td>ev2</td>
<td></td>
</tr>
<tr>
<td>OK?</td>
<td>evaluate</td>
<td>ev2</td>
<td></td>
</tr>
<tr>
<td>Adapt Item</td>
<td>correct</td>
<td>cr2</td>
<td></td>
</tr>
</tbody>
</table>

In Table 6.3, the proposed codes related to the participants’ level of confidence are listed. The assignment of these codes is based on the level of confidence that the participant expresses during the completion of the task. These ‘expressions of confidence’ can take two forms. First, the participant can say that he is (not) sure: ‘I am not sure’, ‘I suspect’, ‘I guess’, etc. Second, the level of confidence can be derived from the actual verbal utterances on the videos: intonations, pauses between words, etc. Both types are combined to assign a level of confidence (confident, not confident, or neutral) to each segmented protocol.
Table 6.3: Proposed codes related to the participants’ level of confidence (Level 3)

<table>
<thead>
<tr>
<th>Confidence</th>
<th>category</th>
<th>code</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confident</td>
<td>confident</td>
<td>cf3</td>
<td>participant is sure that it is correct</td>
</tr>
<tr>
<td>Neutral</td>
<td>neutral</td>
<td>nt3</td>
<td>level of confidence is not expressed</td>
</tr>
<tr>
<td>Not Confident</td>
<td>not confident</td>
<td>nc3</td>
<td>participant is not sure that it is correct</td>
</tr>
</tbody>
</table>

The fourth level (see Table 6.4) is related to the actual actions of the participants during the task. While thinking out loud (talking) the participant is also doing something else to complete the assignment: drawing, checking the drawing, erasing a part of the drawing, taking a pencil, etc. Since these actions occur concurrent with the verbalisations, they have to be coded separately.

Table 6.4: Proposed codes related to the participants’ actions (Level 4)

<table>
<thead>
<tr>
<th>Action</th>
<th>Category</th>
<th>code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Draw</td>
<td>Draw</td>
<td>dw4</td>
<td>participant draws an item</td>
</tr>
<tr>
<td>Take Pencil</td>
<td>take pencil</td>
<td>tp4</td>
<td>participant takes a pencil</td>
</tr>
<tr>
<td>Correct</td>
<td>Correct</td>
<td>cr4</td>
<td>participant corrects item that was drawn previously</td>
</tr>
<tr>
<td>Erase</td>
<td>Erase</td>
<td>er4</td>
<td>participant uses the eraser</td>
</tr>
<tr>
<td>Fill Colour</td>
<td>fill colour</td>
<td>fc4</td>
<td>participant is filling up items with colour</td>
</tr>
<tr>
<td>Checking</td>
<td>Checking</td>
<td>ck4</td>
<td>participant checks drawing</td>
</tr>
<tr>
<td>Talking</td>
<td>Talking</td>
<td>tk4</td>
<td>participant is only talking</td>
</tr>
</tbody>
</table>

In contrast to previous tables, the codes listed in Table 6.5 will not be assigned to the sentence segmentations. Multiple codes would have to be assigned to a single segmented protocol in this case. These codes are used to categorise the aggregated word segmentations. The number of occurrences of each word is counted, and a code will be assigned to a subset of them. This type of categorisation is necessary to further aggregate the data, facilitating comparison operations between the two user groups.

Table 6.5: Proposed codes related to the word segmentation (Level 5)

<table>
<thead>
<tr>
<th>word categories</th>
<th>Category</th>
<th>code</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colour</td>
<td>Colour</td>
<td>cl5</td>
<td>object is described by colour</td>
</tr>
<tr>
<td>Geometry</td>
<td>Geometry</td>
<td>gm5</td>
<td>object is described by its geometry (size, shape,...)</td>
</tr>
<tr>
<td>Name Object</td>
<td>name object</td>
<td>no5</td>
<td>object is recognised and named (forest, city,...)</td>
</tr>
<tr>
<td>Location</td>
<td>location</td>
<td>lc5</td>
<td>object's location is described (left, right, above,...)</td>
</tr>
<tr>
<td>Strange Item</td>
<td>strange item</td>
<td>si5</td>
<td>strange item is mentioned in the description of the object (turtle, ghost, Africa,...)</td>
</tr>
</tbody>
</table>
The categories listed in Table 6.5 are based on how the visual information is encoded in the users’ memory. As mentioned in Section 6.1.1, visual images of maps can be described by concepts or features (their shape, shape, location, colour, etc.) and structural information (spatial framework, distances, directions, relations, etc.) (Kulhavy & Stock, 1996; Thorndyke & Stasz, 1980). The codes in Table 6.5 focus on how the features are described. The structural information can be retrieved from the context in which the individual words are used.

6.2.4.4. The coded protocols
To obtain the coded protocols, the segmented protocols have to be linked to the proposed codes in the scheme. The resulting coded protocols (of the sentence segmentation) will have four codes assigned to each protocol: map level, item level, level of confidence, and action level. This process of assigning codes to protocols was accomplished using the software ELAN. The (sentence) segmented protocols can be imported and linked to the two videos (from the webcam and HandyCam). Using the Media Synchronisation Modus in ELAN, these two videos (and the included audio) can be manually synchronised, based on the placement of the coloured post-its. After the synchronisation step, both videos can be played simultaneously and linked to each protocol.

The proposed codes can be imported in the software using templates for Controlled Vocabularies and Tiers. Finally, in the Transcription Modus of the program, the codes of all four levels can be assigned to each of the segmented protocols in an efficient way. In order to maintain objectivity during the coding process, a subset of the data was coded by three independent coders and their results were compared with these of the main coder. A correspondence of 75.5% between the main coder and the three independent coders was reached, which was considered acceptable.

The coded protocols derived from the word segmentation were obtained differently. The codes listed in Table 6.5 were used to categorise a subset of the (aggregated) list with words and their number of occurrences. The results that are obtained from the analyses of both segmentations are described in Section 6.3.
6.3. Results

6.3.1. Word segmentation

The aggregated word segmentation lists for each (aggregated) word, 1,189 in total, its number of occurrences in the verbal protocols. However, these aggregated data show considerable differences in the total word count between expert and novice participants, with 21,332 versus 13,747 respectively. This difference could be explained by the fact that no fixed time limit was set on the completion of this part of the assignment. Experts took generally longer to complete it. In order to be able to compare the use of wording between experts and novices in an objective way, a normalised set of values have to be used. These normalised values are calculated by setting the total word count of each group equal to 1,000. The occurrence of each word thus corresponds to its ratio in comparison to the total word count of 1,000 (in %).

The total number of occurrences for each word (sum of experts and novices) are ordered from largest to smallest. The word with the highest total number of occurrences is ‘een’ (‘a’ in English), with a ratio of 97.42‰ and 89.44‰ for the expert and novice participants respectively. These ratios are compared between the expert and novice group for each word individually: \((\text{exp}‰ - \text{nov}‰)/ (\text{exp}‰ + \text{nov}‰)\). The result is a value between -1 (only the novices used the word) and 1 (only the experts used the word). In the case of the word ‘een’ this values is 0.043, which indicates that the experts used it only slightly more than the novices.

Next, these test values (TVs) are classified to indicate the level of difference in frequency of word use between expert and novice users. The first class contains the test value (and thus the related words), which show very little difference between the expert an novice users \([0.00;0.25]\). In the second class, this test value is between 0.25 and 0.50, indicating a more pronounced difference. The third class contains words that have an even larger deviating frequency of use between both user groups, with a test value in the interval \([0.50;1.0]\). Finally, the last class contains the words that were only used by one group of users. These categories are thus based on the absolute test values. However, the sign of these test values indicates which user group used the word more frequently: a negative result indicates that the novices
used the word more frequently; a positive result indicates that the experts used the word more frequently.

Figure 6.7 depicts an overview of the occurrence of these test values for expert and novice users across the first half of the ordered word list (first 50% or 595 words). The first group on the left side of the graph depicts 10% of the words that are most frequently used; the second group corresponds to the next 10% of words in the list, etc. The left bar in each group visualises all positive test values (experts use them most frequently); the right bar all negative test values (novices use them most frequently). The sum of both graphs always corresponds to 119 words (10% of the list with words). The shades of gray indicate the amount of difference with the other user group, based on the TVs’ classes.

A trend can be noticed in the bar chart in Figure 6.7. The most occurring words (first 20 to 30%) show a very high similarity between the experts and novices: very high light grey bars. The height of these light grey bars (associated with a very small difference of less than 0.25) decrease systematically until the group of 30-40% of the data is reached. The height of the bars associated with the second category ($TV$ between 0.25 and 0.50) is relatively stable. However, the bar heights of the highest categories ($TV$ between 0.50 to 1) show an increase over the first three groups. Furthermore, the highest category (which indicates that only one user group used a certain word) shows a considerable increase over all groups. It can thus be concluded that the words that are most frequently used are very similar between expert and novice users. However, a diverging trend is noticed between experts and novices regarding words that are less frequently used. The less frequently used words are thus influenced by the users’ level of expertise.
When studying the list of 119 words that were most frequently used (first 10%), it can be noticed that these are all very general, commonly used words: ‘tekenen’, ‘ook’, ‘dan’, ‘in’, ‘links’, ‘huis’, ‘kant’, etc. (in English: ‘drawing’, ‘also’, ‘than’, ‘in’, ‘left’, ‘house’, ‘side’). Words that are more frequently used by the novice users ($TV >0.50$) seem to correspond to general descriptive terms such as ‘green’, ‘middle’, ‘red’, ‘line’, ‘area’. In the second group of the ordered word list, the distinction in the use of specific wording is starting to show. The words that were used (considerably) more frequently ($TV >0.50$) by the novices correspond again to general descriptive terms, including colours: ‘orange’, ‘part’, ‘surface’, ‘both’, ‘beige’, etc. The words that were more frequently used by the experts seem to reflect their level of expertise: ‘north’, ‘field’, ‘core’, ‘south’, ‘meander’, etc.

This initial trend is also visible in the subsequent groups in the ordered list. Experts seem to be able to name the objects using the exact terms, such as ‘farm’, ‘drainage-basin’, ‘cluster’, ‘industry’, ‘agglomeration’, ‘quarry’, ‘electricity wire’, ‘marshland’, etc. The novices use more general words to describe the shape of the objects, including comparisons with other objects that seems strange in this context, such as ‘turtle’, ‘ghost’ or ‘Africa’. Furthermore, novices seem to use relative object locations more frequently (‘top right’, ‘below left’, ‘inside’, ‘straight ahead’, etc.).
This is in contrast with the experts who use wordings like ‘north’, or ‘southwest’ to describe locations. These differences in word use to describe the map content are more pronounced near and in the second half of the ordered list of words.

In Figure 6.8, the word segmentations are grouped based on the codes presented in Table 6.5: the expression of colour, geometry, locations, the name of objects, and strange objects. Likewise as in Figure 6.7, a comparison is made regarding the frequency with which expert and novice participants used the words in these thematic categories. The calculation and classification of these test values (TVs), to analyse which words are more frequently used by experts or novices, is done in the same way as with the ordered word list. The two most striking differences in Figure 8 are to be found in the categories ‘geometry’ and ‘object’. The novice users show a tendency to describe objects based on their geometry: ‘line’, ‘surface’, ‘thin’, ‘curve’, ‘bloc’, etc. This is also found in the verbalisations of the experts, but less pronounced.

Figure 6.8: Overview of the word segmentation, based on a thematic categorisation

The experts tend to use the actual name of the object instead of giving the description of its shape. This is reflected in the two bars related to the thematic
category ‘object’. The experts’ bar is much higher, with a large section depicted in dark grey colours. This indicates that experts use these words much more frequently than novices. Even when studying the actual words that the novices used more frequently than the experts, it should be noted that these are less specific or detailed objects. The novices, on the one hand, used for example ‘plain’, ‘crossroad’, ‘ground’, ‘parcel’, ‘residential area’, ‘field’, ‘hill’, ‘sand’, ‘dirt track’, etc. The experts, on the other hand, named more detailed objects such as ‘drainage-basin’, ‘meander’, ‘bridge’, ‘farm’, ‘quarry’, ‘coniferous’, ‘electric wire’, ‘marshland’, ‘hamlet’, ‘firebreak’, etc.

Only two colours were used considerably more by the expert participants: purple and pink. These colours correspond to less obvious objects in the map, such as the electrical wires. Furthermore, the large portion of dark grey in the experts’ bar related to location can be explained by the use of compass directions such as ‘north’, ‘northwest’ or ‘west’. These were almost never used by novice users. Finally, the thematic groups that contains the ‘strange words’ is more pronounced with the novice users, which can also be explained by the fact that they try to describe the shape of the objects instead of naming them. They use ‘strange objects’ with a similar shape to describe this. The words ‘Asia’, ‘Africa’, and ‘turtle’ were for example used by multiple persons to describe the shape of different forest patches. Experts also used these comparisons with other objects to describe the shape of objects, but less frequently. Strange words only used by the expert group were for example ‘candlestick’ and ‘catapult’ to describe the shape of the main roads.

6.3.2. Sentence segmentation: coded protocols
Besides the word segmentations, the verbal protocols were also segmented based on the criteria of ‘full thought’. This coarser segmentation granularity allows taking the contexts of the individual words into account. Furthermore, one segment (a full thought) can more easily be linked to a specific cognitive process. A psychological model and coding scheme was proposed in a previous section (see Section 6.2.4). Each segment is linked to a cognitive process using this coding scheme, resulting in coded protocols. When analyzing these coded protocols, the time ratio of each segment (or code) in comparison to the whole trial duration, is taken into account. This time ratio is summed for all corresponding codes, taking into account the four
different stimuli and two user groups. The results of these coded protocols are presented in Figure 6.9. The total sum of all bar heights from one user group thus corresponds to 1 within each graph. For example for map 1 the experts’ results for each separate code in Level 1 are ev1=.219, ex1=.754, and or1=.027. The summed time ratios of these codes should thus cover the total trial duration, which is equal to one.

A first glance at these graphs shows that the general trend – the time ratios for each code or cognitive process – is similar for both user groups; no strong deviations can be noticed. What is more, these general trends are repeated over the four different stimuli (or maps) for each level. This means that the ratio of the different cognitive processes that are addressed to solve the problem at hand (remembering and drawing the maps) are similar for both user groups and independent of the actual...
map content (different topographic maps in this case). At the first level, the shortest time interval is associated with the orientation process, then the evaluation of the results, and then the execution of the task. This latter process takes considerably more time than the others (about 80%).

On the second level, the test persons spent most time gathering their thoughts (around 60%). This corresponds to the consultancy of their WM to check whether suitable information is already present. If not, this information is retrieved from the LTM. Next, about 20% of the time was spent on the drawing task. On this level (Level 2), the code \textit{dw2} was noted down when the participant was not verbalizing his thought, only drawing or taking a pencil. This pause in the verbalisations during the thinking aloud assignment can be explained by the interference of two tasks in the WM. The drawing task consumes a lot of the cognitive load available in the WM, and as a consequence the participant cannot verbalise his thoughts anymore. However, the interference between the drawing and thinking aloud task was not always as strong that the participant stopped talking every time he drew something. The graphs related to Level 4 (which list all actions, independent of the verbalisations) show that the participants were drawing about 50-60% of the time and taking a pencil during about 10% of the time, which is much more than the 20% that was registered on Level 2. The participants spent about 5-10% of their time evaluating the objects which they had just drawn, and only 1-2% on the correction of these evaluated objects.

While drawing, the participants could express their level of confidence by using specific words (I think, I guess, I am sure, etc.) or by intonations and hesitations. Most of the time, no clear expression of confidence was detected (\textit{nt3}). The expressions of confidence versus not confidence seem to be in balance for all four stimuli (10-20% of the time). The actions that the participants performed during the verbalisations were recorded on Level 4. As mentioned before, the participants spent most of their time drawing, which is followed by checking (\textit{ck4}). This checking corresponds to evaluating the map (is everything on it, do I miss something, counting the number of objects on the map, etc.) and evaluating individual objects (is the shape correct, is the location correct, etc.). The smallest intervals are related to the correction of items (\textit{cr4}), erasing items (\textit{er4}; which is a part of correcting on
the second level, \( cr2 \), filling the objects’ interior with colours \( fc4 \), and talking without any other operation \( tk4 \).

A more detailed look on these graphs allows identifying smaller variations in the cognitive behaviour of expert and novice map users. The experts, for example, spent more time evaluating the result, whereas the novices dedicate more time to the execution of the task. These differences between expert and novices users tend to become smaller over each trial. On the second level, the experts seem to devote more time on gathering their thoughts \( gt2 \) and evaluating their objects \( ev2 \) than the novices, who spent more time drawing \( dw2 \). Furthermore, during the first trial expert participants expressed their confidence (both higher and lower) more than novice participants. However, during the last two trials, the expressed level of confidence is very similar. The experts’ tendency to check and evaluate their results more than the novices, (see Level 1 and Level 2) is also reflected in the participants’ actions \( ck4 \). Nevertheless, the bar heights related to code \( dw4 \) (drawing) shows a variation between the expert and novice users. In the first trial, the novices use considerably more of the trial time to draw, which is nearly equal between both user groups for the second trial. For the third and the fourth trial, however, the experts spend more time drawing.

6.3.3. **Sentence segmentation: contexts**

While analyzing the segmented protocols based on the full thought, it came apparent that experts also use their knowledge to place objects on a logical position, although they might not even remember that object. This deductive reasoning enables them to solve the assignment more efficiently. For example, if experts draw a village they know it should connected to at least one road and then they try to find a logical location for this road. If there is a valley with a river in it, the experts find it logical that there is also a road and a railway located next to this river. Furthermore, experts try to figure out the reason behind the structure of the map: the direction of the rivers’ flow, elevations, physical properties of the soil, etc.

The novice users do not have the same geographical background knowledge, and are less tempted to use deduction to place an object. They would guess, place an object on a certain location because that part of the map was still rather empty, but they do
not give a specific (valid) reason for doing so. As mentioned before, novice users tend to describe the objects instead of naming them. However, when the novices do name an object, it was noticed during the analyses, that they often use a wrong name. They, for example, said that they were drawing a railway, whereas this was in fact not the case. This ‘railway’ was presented by a thick red line on the screen map, which actually corresponds to a major road. The novices also talked about ‘sandy areas’, whereas these were in fact fields (depicted in yellow).

Some of the colours in the second map were adapted in order to study the influence of this deviating colour use on the participants’ cognitive processes. The colour of the water bodies was changed from cyan to light orange and the background of the villages from light yellow to purple. Most of the experts said that the second stimulus had objects with a natural shape, similar like a meander (top left object) or a drainage-basin (top right object). Some said that it would have been a river if it was blue, but considering its colour might be some kind of quarry. Most of the novices just described it as a capricious beige spot, looking like a worm. Two of them mentioned that it could have been a river if it was depicted in blue. No influence was found regarding the deviating colour use for the village backgrounds.

The mirror operation between the first and last stimuli was also mentioned by several participants at the beginning of the fourth trial. The participants said that they recognised the objects and suspected that the fourth map was a mirrored version of the first one. However, several participants also indicated that this did not facilitate the task: they recognised the objects but they were on a different position. They said that it was difficult to distinguish between the objects in the first and fourth map.

6.3.4. Score on drawn maps
The final maps that were drawn by the participants were compared with the original screen maps. Based on how well these maps were drawn, a score was assigned to them. The criteria ‘how well’ was mainly based on the question, ‘is the object present on the map?’. The aesthetics of the sketched objects or drawing skill of the participants were not taken into account. A similar scoring system as used by Kulhavy and Stock (1996) and Thorndyke and Stasz (1980) was applied to obtain a
score for each map. If the object was present and drawn on the correct location a score of one point was assigned to it. If the object was present, but drawn on a considerably wrong location, only half a point was assigned to it.

The objects that were drawn on the four maps contained villages, rivers, lakes, farms, roads, fields, forests, railroads, etc. In map 1 and map 4, 34 objects were checked, 27 in map 2, and 32 in map 3. These fit in the 19 categories of objects that are typically drawn on sketch maps, as identified by Blaser (2000). These 19 categories are aggregated into five major themes: hydrography, land cover, settlements, roads, and other. The class ‘other’ contained some detailed objects such as electrical wires and islands in the river. The obtained scores were recalculated to percentages (%). These scores are presented in Figure 6.10, on the left side for each theme (all maps) and on the right side for each map independently.

From these graphs it can be derived that the expert participants score considerably better on all categories than the novices. The most pronounced difference can be found in the theme ‘other’, which contained less obvious objects on the map. Both user groups obtain the best scores in the theme hydrography. This seems to correspond to the findings regarding the order with which the objects were drawn. These hydrographic objects were drawn more frequently in the beginning of the assignment, resulting in more accurate results. Next are the settlements and the roads. These roads were also frequently drawn in the beginning of the assignment.
The lowest overall score was obtained for the first map, both by the experts and novices. This could be explained by the fact that this was the first time they had to execute this assignment. Most participants seem to have underestimated the assignment and did not study the map good or long enough. The second map, however, corresponds to the highest score: the participants seem to have studied the stimulus better than the initial one. The fourth map shows again an increase of the score in comparison to the third map. However, this higher score is not as pronounced as could be expected. This fourth map is the mirrored equivalent of the first map, and contains consequently the same objects. Since the participants saw these objects for the second time (however mirrored) it could be expected that they would have remembered this in more detail. As can be derived from the bar heights, the participants’ performance is not significantly better. During this trial they seemed to be confused by the mirrored appearance of these objects.

6.3.5. Order of drawing

The order with which the different structuring elements on the map were drawn can give insights in how these elements were stored in the users’ memory. Elements that are systematically among the first objects that are drawn are more accessible in the LTM; they can be retrieved more easily. By studying patterns in the order with which objects are drawn, it can be derived which map elements are considered more important, and are retrieved more easily. In their article, Huynh et al. (2008) also investigate the order with which elements are recalled while drawing a sketch map. They argue that this sequence gives insights in the hierarchical structure of the users’ cognitive map: a basic framework in which lower level elements are subsequently placed. This hierarchical structure is closely linked with the perceived importance of objects by the user (Huynh & Doherty, 2007; Huynh, et al. 2008).

For each map, eight categories of map elements are identified. These categories contain structuring map elements and are an extension of the four categories used to present the map scores: electrical wire (other), field (landcover), forest (landcover), settlement, hydrography, meadow (landcover), railroad (other), and road. A ‘score’ was assigned to each category based on the order with which the first element of that category was drawn: 100 for the first element, 50 for the second, 25 for the third, 5 for the fourth, 4 for the fifth, and so on. If a certain element was not
drawn, no score was assigned. The graphs visualizing the experts and novices their results are depicted in Figure 6.11, for each map separately. The three categories that are associated with the highest scores are the forests, roads and hydrography. These form the main structuring elements on the maps.

Figure 6.11: The order with which the objects were drawn for each map

In the first map, subjects first drew the forests, the hydrography, and the roads; however a different order for the experts and novices can be noticed. Most novices drew first the forest, than the hydrographic elements, and then the roads, whereas the experts mainly chose to draw the hydrographic elements at first. Less difference was detected between the draw order of the roads and forests for this user group. The same structure of these bars (their relative heights) is also present in the graph related to map 4. This last stimulus was the mirrored version of the first one, which means that the same elements were present. However, the bar that represents the score assigned to the hydrography is not as high, especially not for the expert groups. The novices’ bar related to the forests is also much lower. Nevertheless, both user groups tended to draw the roads earlier than in the first map. Finally, the
detailed objects also received a higher ‘score’ in the fourth map. This indicates that more participants remembered these objects when the map object (although mirrored) were displayed for the second time.

In the second map, the roads were drawn most frequently as initial objects. This is followed by habitation (experts) or hydrography (novices). The colour of the hydrographic objects was adapted in this map, which could explain the lower score in comparison to the other maps. In the third map, the experts drew the hydrographic elements most frequently as the initial objects, whereas the novices focused more on the roads. All other objects received a much lower score. In all maps, the objects ‘railroad’ and ‘electrical wire’ were never drawn as first element. However, from the graphs it could be derived that more experts drew these element than novices. A more detailed analyses on ‘how good’ the maps were drawn can be found in the next section (see Section 6.3.6)

6.3.6. Questionnaire results
After the participants had completed the four trials, they had to fill in a questionnaire. In this questionnaire, they had to report personal characteristics such as age, gender, level of education, and so on. They also had to indicate if the recognised any of the regions that were displayed during the study, as this could bias the results. None of the 24 participants indicated that they recognised one of the regions. They were also asked if they worked with Belgian topographic maps before, how often, and in what context. On this question, the experts noted down that they did use this type of map before during their studies and sometimes still do for their work. Most novices had not worked with these maps before, or not frequently. Furthermore, the participants also had to indicate if they noticed the deviating colour use on the second stimulus and that the fourth stimulus was a mirrored version of the first one. On this latter question, most participants (22 out of 24) said that they did notice it. However, few participants seemed to have noticed the adapted colour use on the second stimuli: 6 out of 24, from which four were experts.

Finally, the participants also had to indicate how well they thought they had drawn each of the maps on a five point scale (ranging from ‘very bad’ to ‘very well’). A score was assigned to each option (-2|-1|0|1|2) and these were averaged for each
map, separately for the two user groups. These average scores of confidence are presented in Figure 6.12. No clear trend can be noticed between the expert and novice users. Furthermore, the level of confidence regarding ‘how well’ the maps are drawn does not agree between the two user groups.

![Figure 6.12: Result of the questionnaire regarding the confidence of the result (drawn map)](image)

The experts are most confident regarding the second and the third map, whereas the novices seem to appreciate their results for the first and last map more. Only the experts have an overall negative score for one map, which is the first one. This was the initial map and the experts seem to be most displeased with these first drawn maps. The score of their fourth map is equal to zero, indicating that they were not pleased with the result, but also not displeased. The novices give a better score for their fourth map in comparison with the experts, although this is also not considered very good. Both user groups seem to be most confident regarding the result of the third map, which is more pronounced in the expert group. This, however, is also not reflected in the actual map scores (see Figure 6.10). The novices are most displeased with their results regarding the second map.

### 6.4. Discussion

The thinking aloud method was used to study the map users’ cognitive processes during a drawing task. The participants had to retrieve information regarding a screen map they had just studied from their LTM. The verbalisations and derived
segmented protocols can be linked to the different cognitive processes that are addressed during the drawing assignment. These protocols give direct, unfiltered insights in the information contained in the WM, which has to be retrieved from the LTM. Especially the difference between expert and novice map users is considered in this.

The analyses of the segmented protocols (full thoughts and word counts) indicate that the structure of the users’ cognitive processes during this task is very similar. What is more, both user groups first draw certain main structuring elements on the maps, a reference frame, and then the detailed objects. This reference frame is generally based on the location of the main rivers or roads in the map. In their experiments, Huynh and Doherty (2007) and Huynh, et al. (2008) obtained similar results. They argued that the sequence of drawing (and thus the recalling from memory) corresponds to the hierarchical structures of the schemas stored in the LTM. They found that their participants tended to draw first the main linear structures and only later on the landmarks. They concluded that these cognitive maps are thus hybrid (consisting out of both paths and nodes) and hierarchically structured.

However, the experts’ results (the maps that were drawn) contain more, and more detailed objects than the novices. This better map score can, on the one hand, be explained by a higher efficiency in the interpretation of maps (see Chapter 2 and Chapter 5). In these chapters the eye movements of experts and novices who studied a screen map with a very basic design (Chapter 2) and topographic screen maps (Chapter 5) were analysed. It could be concluded that the experts can interpret a larger part of the map in the same amount of time as novices. As a consequence, experts can store more detailed information regarding these maps. Furthermore, the experts seem less distracted by deviating elements on the map, such as strange colours or mirrored objects.

The analysed protocols also indicate, on the other hand, that experts can retrieve the stored information regarding a map they had studied more efficiently. They know the name of specific objects on the map, which can serve as a pointer to schemas in the LTM that contain more extensive information regarding these objects. These
pointers facilitate information retrieval from the LTM: the object itself and all related background knowledge. This information comes in the form of a tight group that can form one chunk of information in the WM. Furthermore, this background information can be used in deductive reasoning so that other information (for example, the logical position of linked objects) can be derived from it. This means that experts need to store less information in the LTM because they can retrieve it from knowledge that was already stored.

The background knowledge of the experts is not only related to map use and cartographic syntax, but also regarding geography. In order to be able to interpret a map correctly, the cartographic expertise is of most importance. However, when storing and retrieving information this geographical background knowledge has also a major impact in the users’ cognitive processes: structures in the landscape (past situations, evolutions), physical properties of certain sub soils (vegetation that can grow on it, locations of quarries), economical characteristics (typical locations of industry, structure of road networks), etc. The experts that took part in the experiment have at least a Master degree in Geography or Geomatics and have considerable knowledge regarding these geographical issues, besides the cartographic expertise. While investigating the Knowledge-Is-Power hypothesis, Hambrick and Engle (2002) also demonstrated this facilitative effect that domain knowledge can have on the performance of the users’ memory.

The novices only store and retrieve the information regarding the objects they saw, most of the times without linking it to other information: ‘a red line going from top to bottom over the map’; ‘in the upper part of the map there was a large dark green spot’; ‘there were several yellow spots distributed across the map’; etc. This information is thus stored separately from other information in the LTM, which makes it more difficult to retrieve it. The information that can be retrieved is also less: a description of the shape, the colour and its position (possibly relative to another object on the map). Due to the lack of additional information in the WM, the novices cannot use deductive reasoning to retrieve the position of objects that they did not remember. The novices, for example, need to remember that there was a thick red line (road) and its position in order to be able to draw it, whereas the
experts know (based on deduction) that there has to be a road between two large villages.

Thorndyke and Stasz (1980) and Gilhooly, et al. (1988) did not find any difference in the number of items that experts and novice could draw from memory when considering planimetric maps or de non contour objects of contour maps. They found that the novices spent more time retrieving names, whereas the experts retrieved more information through both lower level and specialised schemas. The stimuli used in the current study were topographic maps on which no names and contour lines could be found. However, the content of the map was rather complex, with many different object types. This complex content could explain why a difference was found in the number of objects that were drawn on the sketch maps by expert and novice participants.

6.5. Conclusion and future work

It can thus be concluded that the experts can retrieve the information from the LTM more efficiently, which also results in the retrieval of more information. Additional background information regarding the objects depicted on the map is also retrieved and placed in the WM. This additional information allows the expert users to use deductive reasoning to derive extra information regarding the map objects. As a consequence, the experts do not have to store this ‘extra information’, in contrast to the novices who cannot use deductive reasoning simply because they do not own as much additional information regarding the objects. Since they cannot derive the information, the novices have to store it as such, which consumes more space in the WM. The additional background information that experts have is contained in the schemas and thus in the (larger) chunks of information that are transferred to the WM.

These findings correspond to what was described by Gobet and Simon (1996, 1998) regarding the expertise in chess, and which was also found in other domains (among others, bridging and computer programming) (Gilhooly, et al., 1998). Experts can access more schemas in the LTM based on the higher amount of background knowledge. This information can be grouped to form larger tight chunks of
information that can be retrieved and placed in the WM. Hambrick and Engle (2002) argued that domain knowledge can function as a retrieval structure which enables the retrieval of specific domain-related information.

A potential bias in this study is that the amount of information that can be retrieved by a certain user is determined by how well the user could interpret the initial visual information. In this case, the expert group also have the advantage (see Chapter 5). However, the strategies that are used to retrieve the information are independent of the interpretation phase, but are influenced more by the background information that was already stored in memory.

These insights in the map users’ cognitive structures, and the differences due to expertise, are an essential contribution to obtain a better understanding of the actual end user of the fast evolving digital cartographic products. Especially the cognitive limits of the novice map users and the differences with the experts are crucial to be able to keep creating effective cartographic products in the future. These products are often characterised by possibilities for user interactions and the display of animations. These elements were not considered in the current study, but are important elements that could have a major influence on the map users’ cognitive processes (and their processing capabilities). These elements will also have to be taken into account in future research.

Acknowledgements
We would like to thank the Belgian national mapping agency (NGI, ‘Nationaal Geografisch Instituut’) to make a number of digital maps from the Belgian topographic map series on 1 : 10 000 available to us, and thus making it possible to incorporate them in the user study.

References


Kulhavy, R. W., & Stock, W. A. (1996). How cognitive maps are learned and

design*. New York: Guilford Press.

*Cartography and Geographic Information Science, 28*(1), 3-12.


Miller, A. M. (1956). The magical number seven, plus or minus two: Some limits on

topographic maps and natural terrain: Effects of experience and task

Montello, D. R. (2009). Cognitive research in GIScience: Recent achievements and

wayfinding: Presentation modes and individual differences. *Spatial Cognition


Nielsen, J. (1994). Estimating the number of subjects needed for a thinking aloud

textual information on navigation in a desktop virtual environment. *Spatial


Slocum, T. A., Blok, C., Jiang, B., Koussoulakou, A., Montello, D. R., Fuhrman, S.,
et al. (2001). Cognitive and usability issues in geovisualisation. *Cartography
and Geographic Information Science, 28*(1), 61-75.


Chapter 7

General Discussion

This chapter summarises and discusses the results obtained throughout the dissertation. Within each of the previous chapters, a detailed description of the achievements is presented. It is the aim of this chapter to provide a link between these findings and place them in a broader context. In the next section, these findings will be summarised and discussed in the light of the research objectives and questions that were identified in Section 1.2.1. Furthermore, a critical reflection on the methodology used to address these research questions will be presented in Section 7.3. Finally, this chapter concludes with challenges and venues for future work (Section 7.4).
7.1. Summary

Table 7.1 presents an overview of the characteristics and findings of the experiments that were described in the previous chapters. In this section, these findings are placed in the context of the different research objectives and questions of this dissertation. Based on the considerations and issues that were described in Section 1.1, the following three research objectives were identified:

Research Objective 1:

*Improve the effectiveness of (screen) map designs based on the users’ characteristics.*

Research Objective 2:

*Contribute to the understanding of how map users read, interpret, store, and use the presented visual information on screen maps.*

Research Objective 3:

*Investigate the influence of (cartographic) expertise on the map users’ cognitive processes and their limitations while processing the visual information presented on screen maps.*

These three research objectives are addressed in four, more specific, research questions. Each of the chapters in this dissertation focuses on different aspects of one or more research questions (see also Table 7.1). In the next sections, these different contributions are summarised and discussed in relation to each research question.
### Table 7.1: Overview of the dissertations’ content

<table>
<thead>
<tr>
<th>Ch</th>
<th>RQ</th>
<th>Map</th>
<th>Participants</th>
<th>Techniques</th>
<th>Analyses</th>
<th>Highlights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Statistical</td>
<td>✓ experts significantly more efficient than novices</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓ evolution CLT: similar both user groups</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓ simple visual analyses</td>
</tr>
<tr>
<td>Ch 2</td>
<td>RQ 1</td>
<td>Basic</td>
<td>Experts</td>
<td>Eye tracking</td>
<td>Statistical</td>
<td>✓ evolution over time: patterns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Novices</td>
<td>Reaction time, measurement,</td>
<td></td>
<td>✓ individual differences</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Questionnaire</td>
<td></td>
<td>✓ influence of map layout</td>
</tr>
<tr>
<td>Ch 3</td>
<td>RQ 1</td>
<td>Basic</td>
<td>Novices</td>
<td>Eye tracking</td>
<td>Visual</td>
<td>✓ introduction of small deviations: label placement</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Reaction time, measurement,</td>
<td></td>
<td>✓ no influence, consciously (questionnaire)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Questionnaire</td>
<td></td>
<td>✓ no influence, unconsciously (measurements)</td>
</tr>
<tr>
<td>Ch 4</td>
<td>RQ 1</td>
<td>Basic</td>
<td>Novices</td>
<td>Eye tracking, Reaction time,</td>
<td>Statistical</td>
<td>✓ focus general map structure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>measurement, Questionnaire</td>
<td>Visual</td>
<td>✓ attracted to deviations in water bodies</td>
</tr>
<tr>
<td></td>
<td>RQ 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓ confusion by mirrored versions</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓ influence of expertise on both</td>
</tr>
<tr>
<td>Ch 5</td>
<td>RQ 1</td>
<td>Complex</td>
<td>Experts</td>
<td>Eye tracking</td>
<td>Statistical</td>
<td>✓ similar cognitive processes experts &amp; novices</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Novices</td>
<td>Questionnaire</td>
<td>Visual</td>
<td>✓ objects are differently stored</td>
</tr>
<tr>
<td></td>
<td>RQ 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓ experts: deductive reasoning, background knowledge, chunking of information</td>
</tr>
<tr>
<td></td>
<td>RQ 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓ novices: descriptive information, need to retrieve all information, can retrieve less</td>
</tr>
</tbody>
</table>
RQ 1: How do map users read and interpret the visual information presented on screen maps?

To understand these processes, a link with existing psychological theories that consider reading and interpretation of visual information has to be established (such as described by among others, Baddeley, 1999; MacEachren, 1995; Montello, 2002).

The reaction time measurements described in Chapter 2 indicate an increase in the users’ cognitive load (as described by Bunch & Lloyd, 2006; Harrower, 2007, Sweller, 1988) while searching for subsequent targets, even on a basic map design. This is a consequence of the limited capacity of the working memory (WM): the number of items it can hold and processing capability (Cowan, 2001; Matlin, 2002). Because of this basic design, the cognitive load needed to process the visual stimuli will be limited (Alvarez & Cavenagh, 2004), which means that a part of the cognitive load for learning (the germane cognitive load) can still be used (Bunch & Lloyd, 2006; Harrower, 2007, Sweller, 1988). After the simulated interaction, shorter reaction times (indicating a lower cognitive load) were registered. Because of the user’s familiarity with a part of the visual stimulus, he had to process and learn less information, resulting in a lower cognitive load according to the model proposed by Sweller (1988).

Studying spatial patterns in the users’ eye movements over time reveals the users’ strategies in their attentive behaviour while trying to interpret the visual stimuli (Matlin, 2002; Wolfe, 1994). A user can only interpret the visual stimulus if his attention is subsequently concentrated on different sub parts of the stimulus. Consequently, this strategy has an influence on how (and which of the) visual information is read and interpreted (i.e., Dodge, et al., 2011; Gobet & Simon, 1996, 1998; Muehrcke, 1986). These visual analyses of the registered eye movements were addressed in Chapters 3-5.

The results related to the basic map designs showed that the users were guided by the location of the labels on the map, starting their visual search
near the list with requested labels. Small individual differences were
registered, but these did not contribute with the same weight as the
influence of the maps’ structure. The patterns resulting from the complex
map designs indicate that the users tended to focus their attention on the
main (linear) structuring elements on the map. The users first tried to store
a general reference frame of map objects. Only later on, more detailed
objects were focussed (and thus processed), so they could be placed in this
reference frame. This is in correspondence with the structure of visual
images as proposed by Kulhavy and Stock (1996), consisting out of feature
and structural information.

Consequently, in order to be able to create more effective maps it is
essential to visualise data in a structured way. Research on change blindness
showed the importance of attention in order to interpret the visual content
of a map (Rensink, 2002). Hegarty, et al. (2010) discovered that map users’
attentive behaviour can be influenced (and thus guided) by the design of the
map (saliency of certain map object). When designing effective a maps (that
facilitate the interpretation process), it is important to stress the image its
general reference frame (such as the major roads and rivers), possibly by
their saliency.

**RQ 2:** How do map users store and retrieve (use) the information that was previously
gathered from screen maps?

Next to the limited amount of chunks that a person can hold in his WM, the
structure of the LTM also plays a key role in how information (derived from
screen maps) is stored in memory (Alvarez & Cavanagh, 2004; Cowan, 2001;
Baddeley, 1999). This structure – with its schemas, links, and pointers –
determines how users can retrieve information stored in the LTM (i.e.,
Atkinson & Shiffrin, 1968; MacEachren, 1995; Matlin, 2002). However,
these theories were not evaluated in the context of information that was
obtained through screen maps. If a map user cannot efficiently retrieve and
utilise the information that was gathered previously from screen maps, there
is no optimal communication process.
By using the thinking aloud method, a direct link to the information stored in the WM can be established (van Someren, *et al.*, 1994). Analysing these verbal utterances on different levels (see Chapter 6) allows studying the cognitive processes during problem solving: retrieval of information and drawing the map in this case (Ericsson, 2006). During the retrieval process, the users tried to establish links between the information in the WM and the information in the LTM. The obtained protocols give insights in what type of links are used to retrieve information stored in schemas. The data indicated that all map users employed the same cognitive processes (ratio, order) in the retrieval process.

Thorndyke and Stasz (1980) and Gilhooly, *et al.* (1988) also studied the different strategies that users employ during information retrieval. These strategies can also have an effect on which and how much information is retrieved. The experiments indicate that this is mainly determined by the user’s background knowledge. This is closely related to RQ 4, the influence of expertise.

These results were also reflected by the complementary data that was collected and analysed, e.g. the order in which the objects were drawn and the scores for the final maps. All users tried to group the information together in order to facilitate the retrieval process. Objects of a same category, for example rivers or forests, were all drawn together as a group. This fits in with the Gestalt Theory that states that humans try to organise or group the object in a visual image (such as maps) (MacEachren, 1995). In this visual image, objects that were part of the map’s main structuring elements were considered most important: they were drawn first, as the users remembered them more clearly.

When combining these results with the discussion of RQ 1, it is clear that the main structuring elements on a map are as important during the interpretation phase, as during the information retrieval phase. In both cases, these map elements function as a reference frame in which other map elements can be placed. This is closely linked to the structure in which these
map elements are stored in (the long term) memory, consisting out of features and structural information (Kulhavy & Stock, 1996). The design of the map objects that contribute to the main map structure is thus an essential in the whole communication process. They should be clearly visible and organised, as such that the map users’ attention guided by them (e.g. Hegarty, et al., 2010; Wolfe, 1994). This facilitated the interpretation phase through which the objects can be stored in memory. If the objects are stored in memory in a structured way, they can also be retrieved more easily (Matlin, 2002).

RQ 3: *How are the map users’ cognitive processes influenced by deviations in the map image?*

Different types of deviations were introduced in the stimuli used during the experiments. These deviations correspond to (less or more pronounced) adaptations in the design (symbolisation) of the map, which are not necessarily conform with the established cartographic rules regarding, for example, colour use (Bertin, 1976; Cuenin, 1972) or label placement (Imhof, 1975).

Research on change blindness (Rensink, 2002; Simons & Ambinder, 2005; among others) has learned that humans are not as capable in detecting changes in a visual image as was assumed previously. The researchers found that map readers could not ‘see’ the change when they were not focussing their attention on the region where the change occurred. In the experiments described in the previous chapters, deviations in the design of certain map elements were introduced. These deviations correspond to changes in the design in comparison to what the participant was familiar with.

In Chapter 4, the aim was to detect whether a small deviation in the label placement (introducing a lower map quality) would influence the effectiveness of the map. The adapted and more striking colour deviation – presented in Chapter 5-6 – was introduced to ensure that a reaction would
be registered. The aim of the experiments was to investigate how the users’ cognitive processes would be influenced by these colour deviations.

When introducing deviations in the structuring elements, the users’ attention was drawn to these items. They were confused by what they registered, because they could not link it to schemas in the LTM (combination of shape and colour of the feature). As a consequence, they could not interpret (or recognise) what they saw using top-down or conceptually driven processing (Gobet & Simon, 1996; Matlin, 2002). What is more, if the map users could not interpret the information, it also could not be stored in the LTM in a structured way: it is not linked to other (familiar) information. Consequently, it was also more difficult to retrieve this information again.

The obtained data indicate also that the interpretation process of the map users was not influenced at all by small deviations, such as the less optimal label placements (see Chapter 4) and different background colour of the villages (see Chapter 5). When studying the attentive behaviour of the map users (see RQ 1) the map users did not focus their attention on these elements, and as a consequence, no difference or deviation was registered. This is in correspondence with the finding related to change blindness (Rensink, 2002).

In this context, the users’ reaction to the mirrored map display (see Chapter 5) could also be seen as a reaction to a deviation, but this time not due to an adapted symbolisation. In this case the users registered a deviation between the current display and something they were already familiar with: the objects were the same, but their location and orientation was different. This initially caused confusion. The users were trying to link the registered information to the information stored in the LTM (the familiar map display), but these two information sources did not agree. This difficulty to learn familiar, though different, elements corresponds to proactive interference (Matlin, 2002). Nevertheless, the familiarity of the map seemed to have facilitated the information retrieval process. A higher score was obtained regarding these drawn maps.
These results indicate that the effectiveness of a map will decrease by a negative influence of top-down processes. Therefore, it should be avoided to significantly alter the design of the structuring elements on a map if the users’ were already familiar with it. However, this is not the case with the design of objects that are not part of the main structures on the map. Their design can be altered (and should not be optimised) as they do not influence the map users’ interpretation process significantly.

**RQ 4:** *How does (cartographic) expertise influence the cognitive processes investigated in the previous research questions?*

As mentioned in Chapter 1, the expertise of the expert group is considered on three levels: cartographic expertise, domain knowledge, and map use. The novice group did not contain any expertise on any of these levels. The level of expertise that someone has could influence his way of thinking and acting, also related to map displays (MacEachren, 1995; Dodge, et al., 2011). Especially in the light of the new cartographic developments and possibilities, it is essential to study both user groups so that guidelines can be identified on how to create effective screen maps for novice map users in the future. As the level of expertise can have an influence on all aspects of the map users’ cognitive processes, this research question is considered in relation to all three previous research questions.

**RQ 1.** From the combined reaction time and eye movement measurements, it could be concluded that novices needed more time to interpret the visual content (simple and complex), resulting in a less efficient visual search. The longer reaction time measurements of the novices also indicated a higher cognitive load, which means that they could utilise less of their cognitive processing capabilities for learning. Similar as with the experts, the novices focussed most of their attention on the main structuring elements on the maps, such as rivers, roads, etc. However, this was still less pronounced than in the expert group.
RQ 2. The results regarding the structures in information retrieval between experts and novices corresponded to the findings of Gobet and Simon (1996, 1998), who studied the cognitive processes of expert chess players. Experts could group their information more tightly together into (large) coherent chunks. This chunking of information is facilitated by the amount of background knowledge that a person has. In this case, experts profited from their cartographic (identifying objects by name) and domain specific knowledge (they have additional information regarding the identified objects). As a consequence, the information in the experts LTM is very structured and possesses many links to background information that was already stored in the LTM. Because of these larger chunks and additional links, the experts could retrieve additional information, allowing them to use deductive reasoning to derive ‘missing’ information.

This is not the case with the novices. They had more difficulties to recognise objects and did not have the additional background knowledge about these items. They only recognised objects on the perception-receipt level, not on the label and other knowledge about level (Gerber, 1981). As a consequence, the recognised objects were stored in an unstructured, descriptive way in the LTM, which made it more difficult to retrieve them. The novices could use less pointers to access the information and the information was not grouped as tightly. Consequently, they needed to retrieve smaller, and thus more, chunks in the WM. Because of the limited number of chunks that can be hold in the WM, this information retrieval process was thus less efficient.

RQ 3. When introducing deviations in the map image, the cognitive processes of both user groups were influenced in a similar way. Where no influence was found regarding small deviations (such as the background colour of villages), the users’ attention was attracted by the more striking deviations. However, this influence was more pronounced in the group of the novices: they focussed longer on these ‘strange’ objects than the experts. This indicates that they needed more time to interpret (recognise) the visual content, by linking it with previous knowledge. A similar effect was observed during the interpretation process of the maps that were depicted ‘upside
down’. In the beginning, all participants seemed confused because they recognised the map objects, which were positioned completely differently. Again, the observed confusion related to these maps was stronger in the novice group than with the experts.

In conclusion, it could be stated that the communication process is less effective on all aspects regarding the novice users: attention, interpretation, information retrieval, and influence by deviations. Especially for this user group it is thus essential to develop a map design that facilitates the interpretation process and, consequently, also the retrieval of information.

7.2. Improving the effectiveness of (screen) maps

Based on the answers on the research questions that were described in the previous section, a number of recommendations can be made to create more effective (screen) maps, taking into account the user’s characteristics (expertise in this case).

The different phases in the communication process (from initial attention to information retrieval) are closely linked. Improvements in one of the initial phases will have a positive effect on all other subsequent phases. Previous research (e.g. Hegarty, et al., 2010; Rensink, 2002; Wolfe, 1994) has indicated that the users’ attentive behaviour can be influenced by altering the design of a visual stimulus (such as a screen map). The focus of the user’s attention determines which elements will be registered in detail through the visual stores and thus interpreted. The information that can be interpreted and subsequently learned is transferred to the user’s (long term) memory. If this is stored in a structured way, it can easily be retrieved again (Matlin, 2002).

As proposed by Hegarty, et al. (2010), a good map design can thus facilitate the interpretation process. Through the design of the elements of the map, the users’ attention is guided to the important information, essential for the interpretation process. In the case of topographic maps, the eye movement data revealed that the main (linear) structuring elements on the map receive most attention (which is somewhat less in the novice group). In order to guide the (novice) map users’
attentive behaviour, these map elements should thus be depicted in more salient (or saturated) colours.

Based on humans’ tendency to organise a visual scene (for example by grouping similar objects) it is also important to aid the map users’ in this process. Object close together and objects that belong to the same category should form clear groups that can be identified through their design. This grouping of information can also be found in the users’ memory structures: interlinking of information in schemas and chunking of information during information retrieval. Clear visualisation of these groups facilitates these processes. However, the design of these structuring map elements should not deviate much from what the participant was already familiar with. This causes a negative effect on the interpretation process, resulting in confusion, because of top-down processing.

Map elements that do not contribute to the main structure of the map are not often in the focus of the user’s attention. As a consequence, their design will not have a significant impact on the interpretation process. When improving the effectiveness of a map, the design of these elements is thus less important. However, care should be taken that these elements are not visualised with striking colours (patterns, sizes, etc.) which would attract the map users’ attention. This would distract them from the structuring elements on the map, which could distort the interpretation process.

The recommendations presented in the previous paragraphs are especially important in the light of recent technological developments. The accessibility of cartographic information for novice map uses has increased significantly through the Internet. Furthermore, with the introduction of Web 2.0, almost all web users can create cartographic products themselves, without considering and applying well established cartographic rules. These amateur maps are then released on the web, were they can be used again by the community.

The novice map makers on the Internet form a powerful source of information that should not be neglected. However, the cartographic products that they create might not be as effective, caused by their lack of cartographic training (Cartwright, 2012). It is not feasible to train all these novice map makers to improve their cartographic
skills. Implementing the wide variety of cartographic rules in the tools that enable them to create web maps could be possible, but this would make them overly complex in use.

The main findings of this PhD regarding to the elements that contribute to an effective communication process through screen maps can be implemented in tools to create web maps. The results indicate, on the one hand, that not all elements on a map have to be designed in the most optimal way (e.g., the elements that do not contribute to the main structure of the map). However, regarding the main structuring elements, a limited number of guidelines or limitations could be implemented, which would improve the effectiveness of these maps significantly.

7.3. Critical reflection on the methodology
The evaluation techniques used in this dissertation are based on the principles used in User Centred Design. Two different stages of the UCD life cycle are implemented in the experiments: early stage evaluation on a basic prototype and evaluation of a final product. In the different experiments, a mixture of techniques was used to study multiple aspects of the research questions and to ensure the validity of the data. The next paragraphs, present a critical reflection on the methods and techniques used to record, analyse, and study the users’ cognitive processes while working with screen maps.

Eye tracking was used to study the map users’ attentive behaviour. The derived eye movement metrics, such as fixation duration and fixation count, can give insights in the users’ cognitive processes. Since eye tracking is not a new technique, already a number of works have discussed the meaning of the obtained metrics in relation to the users’ cognitive processes (i.e., Duchowsky, 2007; Poole & Ball, 2006). Eye tracking has been used in cartographic research in the past, but only recently a renewed interest is noticed (Brodersen, et al., 2001; Çöltekin, et al., 2009, 2010; Fabrikant, et al., 2008). Next to the quantitative analyses, the data were also analysed from a qualitative point of view: visual analysis. This type of analysis allows studying the spatial (and temporal) dimension of the data, including patterns in the users’ scanpaths, in more detail. In the past, eye movements were often visualised
using heatmaps. However, this type of visualisation is not optimally suited to analyse data in a scientific way. In general, objective and detailed qualitative (visual) analyses of eye movement data is still limited in scientific research. In this dissertation, new methods are proposed to extend the possibilities regarding the visual analyses of eye movement data (see Chapter 5). Besides analyses described in this dissertation, other methods have been proposed, such as sequence analyses (Çöltekin, et al., 2010) or visual analysis in the Space Time Cube (Li, et al., 2010).

During the visual search assignment, the eye tracking data was complemented with reaction time measurements (see Chapters 2 and 4). In usability engineering, reaction time measurements are often used to evaluate how fast (performance) a user can accomplish a certain task on a system (Nielsen, 1993). In this dissertation, these reaction time measurements are analysed in combination with the recorded eye movements in order to verify their validity. The location of the participants’ point of regard was identified at the moment of each reaction time measurement and compared to the locations of the target labels. Incorrect measurements (locating a wrong target label) could thus be removed from the dataset in order not to influence the analyses and final outcomes.

In order to study the users’ cognitive processes during information retrieval, the thinking aloud technique was used (see Chapter 6). With this method, a direct link with the users’ thoughts in the WM is established. The verbal utterances can also be analysed in different ways, using different segmentation granularities or approaches (Chi, 1997; van Someren, et al., 1994). Also in this case, it was decided to combine these different approaches (segmentation and coding levels) to analyse multiple aspects of the data and add a validity check to the outcomes of the different analyses.

All experiments concluded with a post study questionnaire. This questionnaire was essential to gather background information regarding the participants: age, gender, education level, familiarity with map use, etc. Furthermore, the participants had to answer questions regarding the experiment itself: familiarity with the displayed region, level of confidence of the result, check whether they noticed certain elements on the map, etc. Based on this data, it could be verified whether the participants’ characteristics complied with the characteristics of one of the two user
groups in the experiments. If a characteristic of a certain participant did not correspond to the criteria that were set (such as the familiarity with a certain region display during the experiments), the related data could be removed from the recordings.

Because of this mixture of techniques and analyses, multiple aspects of the users’ cognitive processes could be analysed, with the incorporation of validity controls. Especially the combination of quantitative and qualitative analyses is an essential aid to investigate the problem at hand from several different viewpoints. As a consequence, it is not only important to know how to apply the methods that are employed in UCD (as investigated by Nivala (2007), but also how to combine them in an effective way to obtain the most complete overview and thus useful results.

7.4. Recommendations for future work
Every dissertation is linked with a number of limitations, due to the decisions that had to be made in every step of the research: structure of the experiments, stimuli, participants, tasks, analyses, etc. Some decisions will not have a significant impact on the final outcomes (such as, ‘Do I provide an eraser during the drawing task?’), whereas others are more critical (such as ‘How many participants do I have to select?’, or ‘Which task do the participants have to complete?’). As a consequence, care has to be taken regarding the applicability of the findings presented in this dissertation on other scenarios: other type of users, stimuli, tasks, etc. However, these limitations can be addressed in future research, which would extend the generalisability of the results. Based on the findings presented in this dissertation, a number of challenges and recommendations for future work are discussed in the next sections, ordered from less to more important.

7.4.1. Future work considering alternative tasks
The tasks which were used in the different experiments were kept rather simple, so that it would not demand a high cognitive load. The visual search assignment is a type of task that can be applied in a wide range of applications in the context of map use. The user is trying to locate some point of interest on the map. This operation often has to be completed before and during a more complex task. Therefore, the
obtained results can be applied to a wide range of other scenarios. However, when users have to perform a more complex or totally different assignment, their eye movement patterns and cognitive processes that are addressed could differ. These tasks could include the analyses of the content of the map in order to derive additional information from it, such as establishing links between different objects. However, the tasks that a user can execute is closely related to the content of the map and thus also to type of the map (reference map, thematical map). Examples of alternative tasks could be: ‘Try to find the ideal location for a new building site, taking into account a number of parameters’, ‘Indicate which administrative unit has the best economic characteristics’, etc.

7.4.2. Future work considering map users
Two user groups were considered in this dissertation: experts and novices. With experts, persons with a high level of cartographic and domain knowledge were selected. Furthermore, the experts were also trained in map interpretation and map use. The novice users did not possess any expertise in any of these three levels. Two extreme groups regarding expertise were thus evaluated in the user tests. If no difference between these two users groups was registered, there would have been no need to test two user groups who are ‘closer’ in terms of expertise. The results in this dissertation indicate, however, that there is a significant difference in how both user groups process the screen maps. As a consequence, it would be useful to include a wider variety of map users, in terms of their expertise, in the tests. This way, it could be investigated what the influence of (only) domain knowledge is in map interpretation and information retrieval, as opposed to expertise in map use or only (theoretical) cartographic expertise. Next to expertise, other user characteristics could be considered – such as gender and age – in order to create more effective maps based on these specific user characteristics.

7.4.3. Future work considering map types
During the experiments, only reference maps were presented to the users. Furthermore, these maps were not overcrowded with objects and had a structured layout. Consequently, the findings described in this dissertation are only applicable to these types of maps. Nevertheless, this type of maps is often used in a wide range of applications, especially on screen displays with a lower resolution.
they are often used as a base map with thematical information superimposed. Next to reference maps, a wide variety of thematical maps exists: choropleth maps, dot maps, cartograms, etc. The specific characteristics of each of these thematical map types could influence the users’ cognitive processes while trying to interpret the content. The implementation of a wider range of map types and more complex designs could thus be considered in future research.

7.4.4. Future work considering user interactions
The evolutions towards and in neocartography have introduced new opportunities in how cartographic information can be represented. In the past, the majority of the digital maps were static and view-only, supporting a unidirectional communication process (Kraak & Brown, 2001). Today, most of the digital cartographic products allow the user to interact with them and as such to manipulate them, both on the level of content and visualisation properties. Users can change the scale (zooming), current extent (panning), amount of information (turning on and off layers), typography (font sizes), etc. These user interactions can also trigger dynamic responses, such as animations which are often used to visualise temporal aspects of the data (i.e., Kraak & Ormeling, 2010; Slocum, et al., 2007).

The findings described in this dissertation are based on controlled experiments, conducted in a laboratory set-up. The users were not allowed to interact with the system to manipulate the visualised stimuli in any way. The controlled nature of the experiments ensured the comparability of the registered data, which facilitates the analyses. The simulated pan operation that was included in the experiments described in Chapters 2-4 showed exactly the same pan operation on all maps and to all users: after the same time interval, in the same direction, and over the same distance. This level of control, and thus comparability of the data, would be lost if the participants could interact with the systems themselves. Especially when using eye tracking as an evaluation method, it is very difficult to analyse data that is linked to unstructured interactions that result in dynamic responses on the screen.

Recently, some work has been undertaken to evaluate interactive map interfaces and animations using eye tracking (i.e., Çöltekin, et al., 2009, 2010; Fabrikant, 2005), but detailed analyses of the data is still a key research challenge. The main challenge
is related to the unstructured changes (both spatially and temporal) in the map and its interface on which the eye movement data are superimposed. As a consequence, it is very difficult to (automatically) link the eye movement data to the ever changing display with which the user was working during the experiment. However, this linkage is essential to be able to objectively evaluate the eye movement data. This specific research challenge can thus be translated in a critical venue for future research. If a framework can be built which would facilitate the analysis of eye movement data linked with interactive and dynamic displays, it would be possible to investigate the map users’ cognitive processes while working with these types of cartographic products.

References


Chapter 8

General Conclusion
This dissertation comprises a methodological contribution to the evaluation of
digital cartographic products and its users, based on User Centred Design (UCD).
Throughout the chapters, a number of techniques that originate from UCD are
discussed: questionnaires, eye tracking, reaction time measurements and thinking
aloud. The studies indicate their suitability to evaluate (digital) cartographic
products, but especially the combination of different techniques has proven to be an
essential asset to obtain the most complete overview of the issues under
investigation.

Besides the different techniques used, special attention was paid to the analyses of
the data, with a focus on both its quantitative (statistical) and qualitative (visual and
spatial) aspects. The results from these in depth analyses show that the different
approaches (qualitative vs. quantitative) complement each other, on the one hand,
and support their validation, on the other hand. This dissertation demonstrated that
combining these type op analyses is crucial in the field of cartography and (by
extension) geovisualisation.

In this dissertation, the focus is directed towards questions regarding how map users
read, interpret, store and retrieve the visual information presented on screen maps.
Special attention is paid to the differences between two distinct user groups: experts
and novices. In general, it can be concluded that a similar structure could be
identified in the users’ attentive behaviour and cognitive processes. The obtained
eye movements reflect the main structuring elements on the map, which indicate
that both user groups focus their attention on this reference frame during the
interpretation process. Furthermore, small deviations (adaptations in the objects’
colour or location) in the map image do not influence the users’ cognitive processes,
whereas more striking deviations attract the users’ attention or cause confusion.
Finally, the map users’ memory structures and how they retrieve the information
from the Long Term Memory (LTM) also corresponds between both user groups, by
linking the information in the Working Memory (WM) to this stored in the LTM.

However, the user studies also showed significant differences in how both user
groups interpreted and retrieved the visual information. The experts’ measurements
demonstrate that these users are more efficient on all accounts. The recorded eye
movements indicate that experts can interpret (and thus process) the visual image faster, which implies that they can also search on a larger part of the map in the same amount of time (compared to the novices). Consequently, they can find a certain point of interest in a shorter time interval than the novices.

What is more, the experts focus their attention on the map’s structuring elements, whereas the novices are more distracted by details on the map. Nevertheless, the experts can remember more detailed elements when retrieving the information from memory. This is a consequence of the more extensive amount of background knowledge that experts have (both cartographic and domain specific). As a consequence, experts can store this information in a more structured way (more interconnected links with the new and old data in the LTM), which also means that they can access and retrieve it more easily. Through the established links with the ‘old information’, the experts can also retrieve additional information that was not directly gathered from the map image (using deductive reasoning).

Novices store almost everything they interpret separately because fewer links with existing (old) information can be established. This implies that more information has to be stored, which is also less structured and thus less accessible. Because novices cannot ‘chunk’ their information as efficiently as the experts, they cannot retrieve (and thus remember) as much information from the LTM. What is more, they cannot use any additional (domain specific) background information to derive additional information regarding the elements on the map.

Finally, both user groups also react differently when they notice deviations (adaptations) in the map image. Regarding small differences (location of labels and colour of less important or structuring objects) no reaction in both user groups is measured. Nevertheless, when more striking deviations are presented, both user groups react differently. Regarding (structuring) objects that are visualised in a ‘strange’ colour, the attention of both user groups is attracted to it. However, the novices focus their attention much longer on these deviating objects than the experts. When considering the mirrored objects, the novices show confusion, which is not so pronounced in the expert group.
The differences mentioned above have to be taken into account when designing effective cartographic products. During this design process, a number of cartographic guidelines are maintained. However, these guidelines do not take the cognitive limits of novice map users into account. The studies described in this dissertation demonstrate that the cognitive processes of experts and novices differ when interpreting maps (and using this information later on). When experts change the design of a map (introducing deviations) from their point of view, this might thus be interpreted differently by novices. What an expert might consider more effective could be distracting for the novice user. It is thus essential that map makers evaluate their new designs with actual end users before their release.

This dissertation aims at contributing to the understanding of how map users process the visual information on screen maps and thus of the cognitive limits of novice map users. This information is essential to create guidelines for the map designers regarding the map users’ cognitive limits in order to be able to keep creating effective (digital) maps in the future. This would make it possible to key the design of the maps to the users’ cognitive processes, which would facilitate the interpretation process.

What is more, with the new cartographic possibilities in the neocartography, novice map users can also – fairly easily – create and distribute cartographic products. These novice map makers often do not have any understanding of ‘how maps work’. They do not possess the necessary cartographic background knowledge to create effective products and do not consider the cognitive limits of their target audience. However, volunteered geographic information and crowdsourcing are nowadays powerful tools to gather and distribute large amounts of data. If we could implement these cartographic and cognitive guidelines in the tools with which novice map users create their maps, they would be more effective and thus usable.

When considering animations and user interactions, the level of expertise could have an even larger influence on the users’ cognitive limits. In order to be able to keep creating effective maps in the future, it is essential not to stop this research here but to extent it so it can be applied to a wider array of cartographic products. This latter is considered a main research challenge for the near future.
Samenvatting (Dutch Summary)

Kaarten worden al eeuwen gebruikt als communicatiemiddel. De cartograaf tracht (een deel van) de wereld rondom ons zo goed mogelijk voor te stellen en dus over te brengen (te communiceren) naar de kaartgebruiker. Deze informatie moet worden gegeneraliseerd en gesymboliseerd om te kunnen worden voorgesteld op een medium (boek, poster, scherm) met bepaalde beperkingen, zoals zijn afmetingen. Deze symbolisatie kan nochtans ook ‘ruis’ in het communicatieproces introduceren, waardoor de boodschap niet optimaal wordt overgebracht (Dodge, et al., 2011; Keates, 1996; Muehrcke, 1986).

Tijdens de tweede helft van de vorige eeuw werden er een aantal theoretische cartografische regels voorgesteld om de ruimtelijke informatie zo effectief mogelijk te communiceren (o.a. door Bertin, 1967; Cuenin, 1972; Imhof, 1975). Met deze regels werd er onder meer getracht de visuele eigenschappen van symbolen (zoals grootte, vorm, kleur, grijswaarde, orientatie en textuur) te linken aan waarnemingseigenschappen: aantal, volgorde, binding en overzicht. Tot op heden worden de meeste van deze regels nog steeds gehandhaafd in het ontwerp van kaarten. Nochtans is er ondertussen nog maar weinig onderzoek gebeurd waarbij deze theoretische regels geëvalueerd worden op hun daadwerkelijke effectiviteit, door middel van gebruikersstudies waarin de eigenlijke eindgebruiker wordt geraadpleegd.

Gedurende de laatste twintig jaar worden meer en meer digitale cartografische producten gemaakt. Het aantal kaarten dat tegenwoordig op het internet geraadpleegd kan worden, overtreft ruim het aantal papieren kaarten dat dagelijks geproduceerd wordt (Peterson, 2003). In 2001 stelde Kraak een ruwe classificatie van deze internetkaarten voor, waarbij hij onderscheid maakte tussen statische en dynamische kaarten enerzijds en tussen ‘alleen kijken’ en aanklikbare kaarten anderzijds. Tegenwoordig zijn de meeste digitale kaarten aanklikbaar, wat een of andere dynamische reactie veroorzaakt, zoals zoomen, pannen, lagen aan- of uitzetten, lettertype vergroten, etc. Met de komst van Web 2.0 zijn de interactiemogelijkheden nog meer toegenomen. Gebruikers kunnen zelf bijdrages
leveren aan de inhoud van de kaarten door middel van vb. crowdsourcing, wat weer reacties van andere gebruikers kan uitlokken.

Deze digitale kaarten, die op een scherm worden afgebeeld, gaan nochtans gepaard met een aantal belangrijke beperkingen. Lage resoluties, beperkte afmetingen en afwijkende kleurweergaves zijn hier enkele van (Peterson, 1995). Doordat de gebruikers steeds meer toegang hebben tot de inhoud en weergavemogelijkheden van de kaarten op het internet, kunnen zij – vrij gemakkelijk – zelf nieuwe cartografische producten creëren en verder verspreiden, zonder dat zij enige cartografische training hebben gehad. Dit laatste kan eveneens een negatief effect hebben op de effectiviteit van deze schermkaarten en dus op het communicatieproces (Cartwright, 2012).

Een tiental jaar gelden uitte een aantal onderzoekers hun bezorgdheid aangaande al deze nieuwe mogelijkheden, meer bepaald de vele mogelijkheden voor gebruikersinteracties met al dan niet dynamische reacties (MacEachren & Kraak, 2001; Slocum, et al., 2001). De vraag rees in welke mate dit een invloed had op hoe deze informatie wordt geïnterpreteerd en of het dus wel voordelig is voor de effectiviteit van de communicatie van ruimtelijke informatie naar de eindgebruikers.

Met de recente evolutie van de digitale cartografische producten kan er een nauwe band met de softwareontwikkeling worden geïdentificeerd. In dit veld is User Centred Design (UCD) geen onbekende methode om de gebruiksvriendelijkheid van de producten toe te evalueren in verschillende fasen van het ontwerp. Hoewel de technieken die tijdens het UCD aangewend worden (zoals vragenlijsten, luidop denken, interviews, focusgroepen, eye tracking, etc.) ook kunnen gebruikt worden om cartografische producten te evalueren, worden ze in de praktijk nog zeer weinig toegepast (Fabrikant & Lobben, 2009; Montello, 2009; Nivala, 2007).

Met dit doctoraat wordt getracht om een antwoord te bieden op de vragen gerelateerd aan hoe gebruikers visuele informatie op schermkaarten verwerken. Deze inzichten zijn essentieel om effectieve – en dus ook hoogkwalitatieve – digitale producten te kunnen blijven produceren in de toekomst. Hiervoor wordt beroep gedaan op enkele technieken uit de UCD om zo eveneens een methodologische
bijdrage te leveren aan het onderzoek aangaande de evaluatie van digitale cartografische producten waarin de eindgebruikers betrokken worden. In dit doctoraatswerk werden zo drie verschillen onderzoeksdoelstellingen geïdentificeerd:

Onderzoeksdoelstelling 1:

*Een verbetering van de effectiviteit van het ontwerp van (scherm)kaarten, gebaseerd op de gebruikers hun karakteristieken.*

Onderzoeksdoelstelling 2:

*Een bijdrage leveren aan de inzichten inzake hoe kaartgebruikers de voorgestelde visuele informatie op schermkaarten lezen, interpreteren, opslaan en gebruiken.*

Onderzoeksdoelstelling 3:

*De invloed van (cartografische) expertise op de cognitieve processen en beperkingen van kaartgebruikers tijdens de verwerking van de visuele informatie voorgesteld op schermkaarten onderzoeken.*

Deze drie onderzoeksdoelstellingen worden vertaald in vier meer concrete onderzoeksvragen (*OV*) welke in het proefschrift besproken zullen worden:

Onderzoeksvraag 1:

*Hoe lezen en interpreteren kaartgebruikers de visuele informatie die is voorgesteld op schermkaarten?*

Onderzoeksvraag 2:

*Hoe wordt de informatie die eerder verkregen werd van schermkaarten door de kaartgebruikers opgeslagen en gebruikt?*

Onderzoeksvraag 3:

*Hoe worden de cognitieve processen van de kaartgebruikers beïnvloed door afwijkingen in de kaart?*

Onderzoeksvraag 4:

*Hoe worden de cognitive processen, welke onderzocht werden in de voorgaande onderzoeksvragen, beïnvloed door (cartografische) expertise?*

De bovenstaande onderzoeksvragen worden behandeld in hoofdstukken 2 tot en met 6 van dit proefschrift. Twee grote onderdelen kunnen hierin onderscheiden worden. In hoofdstukken 2 tot en met 4 wordt een zeer eenvoudig en gecontroleerd kaartontwerp getoond in de gebruikersstudies. Tijdens deze eerste studies moesten
de deelnemers een aantal namen (labels) terugvinden in het kaartbeeld. Voor elke gevonden naam werd een tijdsmeting bijgehouden. Daarnaast werden de oogbewegingen van de deelnemers tijdens deze zoektocht geregistreerd. In elk van deze eerste drie hoofdstukken wordt steeds een combinatie van meerdere onderzoeksfragen behandeld.

Hoofdstuk 2, waarin de resultaten van een studie met twee verschillende gebruikersgroepen worden besproken, focust op OV1 en OV4. De eerste gebruikersgroep bestaat uit experten in de cartografie en de tweede groep bestaat uit deelnemers die geen cartografische training hebben gehad en niet regelmatig werken met cartografische producten. De geregistreerde oogbewegingen (duur van de fixaties en aantal fixaties) en reactietijdmetingen worden statistisch geanalyseerd en vergeleken. Daarnaast wordt een beperkte visuele analyse voorgesteld. Uit de bekomen resultaten blijkt dat de expertengroep de stimuli (met een eenvoudig, gecontroleerd kaartontwerp) efficiënter kan interpreteren: kortere fixatieduur en meer fixaties per seconde. Aangezien de experten in hetzelfde tijdinterval een groter oppervlak van de kaart kunnen interpreteren dan de niet-expert, kunnen ze de gevraagde namen sneller (efficiënter) terugvinden.

In het derde hoofdstuk wordt de nadruk gelegd op de visuele analyse van de geregistreerde oogbewegingen van de niet-expert. Met deze visuele (en dus kwalitatieve) analyse kan het ruimtelijk aspect van de oogbewegingen in acht genomen worden. Aangezien het doel van kaarten de communicatie van ruimtelijke informatie inhoudt, is dit type van analyse een onontbeerlijke aanvulling op de kwantitatieve, statistische verwerking. Met behulp van de Visual Analytics Toolkit kan de immense hoeveelheid – typisch resulterend uit een eyetrackingexperiment – worden gefilterd (gebaseerd op meerdere attributen) en geaggregeerd. Daarnaast kunnen temporele evoluties van de gebruikers hun oogbewegingen worden bekeken, waardoor de identificatie van patronen mogelijk is. Met deze analyses kan men inzicht verkrijgen in hoe deze gebruikersgroep het kaartbeeld interpreteert (OV1).

In hoofdstuk 4 worden twee verschillende stimuli gepresenteerd aan een homogene gebruikersgroep, meer bepaald de niet-expert. Beide type stimuli verschillen in de manier waarop namen (labels) in het kaartbeeld geplaatst zijn (OV1 & OV3).
het eerste kaarttype worden de labels volgens de algemeen aanvaarde cartografische regels geplaatst. In de tweede groep van stimuli wordt een geoptimaliseerd (qua efficiëntie en dus verwerkingstijd) tekstplaatsingsalgoritme aangewend, wat een verminderde cartografische kwaliteit met zich meebrengt. Een verbeterde algoritmische efficiëntie is van belang wanneer er gewerkt wordt met interactieve en dynamische kaarten, aangezien gebruikers na elke interactie niet lang willen wachten totdat het kaartbeeld ingeladen (berekend) is. De studie die beschreven in dit hoofdstuk is beschreven onderzoekt of deze lagere cartografische kwaliteit een significante invloed heeft op het interpretatieprocess van de gebruiker (niet-expert). De resultaten (afkomstig uit zowel de statistische als de visuele analyses) geven aan de gebruikers geen invloed ondervinden van de lagere cartografische kwaliteit, waardoor het algoritme gebruikt kan worden.

In hoofdstukken 5 tot en met 6 worden de voorgaande studies uitgebreid door het gebruik van meer realistische en complexe schermkaarten als stimuli, gebaseerd op de Belgische topografische kaarten op 1 : 10 000. De opdracht van de deelnemers (zowel experten als niet-experten; OV 4) bestond erin om het getoonde kaartbeeld zo goed mogelijk te onthouden. Gedurende deze periode probeert de deelnemer dus de visuele informatie te lezen en interpreteren (OV 1). Tijdens deze taak werden de oogbewegingen van de deelnemers geregistreerd. Wanneer de deelnemers het kaartbeeld voldoende hadden bestudeerd, moesten zij dit tekenen op papier. Hierdoor moeten ze de eerder opgeslagen informatie terug opvragen uit het geheugen (OV 3). Om dit proces te kunnen registreren (en dus inzicht te verkrijgen in de structuur van het geheugen) moesten de deelnemers luidop denken. In de vier getoonde stimuli zijn eveneens een aantal afwijkingen geïntroduceerd: aangepast kleurgebruik voor het water en dorpskernen en spiegeling van het kaartbeeld.

De geregistreerde oogbewegingen van beide gebruikersgroepen worden zowel statistisch als visueel geganalyseerd en vergeleken (hoofdstuk 5). De statistische analyses bevestigen de bevindingen die in het tweede hoofdstuk beschreven werden. De visuele analyses worden uitgebreid door de uitwerking van een fijnmazig rooster waarin de variabelen (fixatieduur en aantal fixaties, zowel het gemiddelde als de maximale waarden voor elke gebruikersgroep) ruimtelijk worden weergegeven met
behulp van een geschikt kleurschema. Daarnaast wordt een 3D-weergave van de gegevens in dit rooster gebruikt om uitschieters te identificeren en deze te kunnen vergelijken tussen beide gebruikersgroepen. Tot slot wordt een statistische analyse tussen de overeenkomstige roostercellen van twee roosters uitgevoerd en vervolgens gevisualiseerd. Deze visuele analyses geven aan dat de gebruikers voornamelijk focussen op de algemene structurerende elementen in het kaartbeeld (zoals wegen en rivieren) en dat dit meer uitgesproken is bij de expertengroep. Beide gebruikersgroepen zijn afgeleid of tonen verwarring door de afwijkingen in het kaartbeeld, wat meer naar voren komt in de groep van de niet-experten.

De analyse van de verbale protocollen (hoofdstuk 6) die resulteren uit de geregistreerde audio- en video-opnames van het luidop denken – zowel op het niveau van de woorden als de ‘volledige gedachte’ – verschaf inzicht in hoe de deelnemers de informatie (eerder verkregen van een schermkaart) terug opvragen. Hierdoor kan er eveneens inzicht verkregen worden in hoe deze informatie gestructureerd is in het geheugen en de verschillen hierin tussen beide gebruikersgroepen. Hieruit blijkt dat de experts veel meer informatie kunnen opvragen, aangezien zij dit kunnen linken aan achtergrondinformatie die al aanwezig was (expertise in cartografie en het vakgebied). Doordat de informatie gestructureerder opgeslagen is in het geheugen kan hij dus makkelijker opgevraagd worden. Bovendien kunnen de experts deductie gebruiken om informatie over ontbrekende elementen af te leiden. Door het gebrek aan achtergrondinformatie moeten de niet-experten alles in hun geheugen opslaan, en dat gebeurt onder een minder compacte en gestructureerde vorm. Alle informatie wordt afzonderlijk opgeslagen zodat het minder gemakkelijk terug opgevraagd kan worden.

Met dit proefschrift wordt getracht een methodologische bijdrage te leveren aangaande de evaluatie van cartografische producten en zijn eindgebruikers, gebruik makende van principes uit het User Centred Design. In de hoofdstukken worden een aantal studies beschreven waarin (een combinatie van) methoden en technieken worden besproken waarmee zowel verschillende gebruikersgroepen (experten vs. niet-experten) en verschillende kaarttypes (cfr. tekstplaatsing) worden geëvalueerd, zoals eyetracking en luidop denken. Daarnaast wordt er veel aandacht
besteed aan de analyse van de bekomen data, zowel kwantitatief (statistisch) als kwalitatief (visueel en ruimtelijk).

Uit de studies is gebleken dat er een significant verschil kan worden waargenomen tussen hoe experten en niet-experten schermkaarten lezen, interpreteren, opslaan in het geheugen en dit vervolgens weer opvragen. Deze verschillen zijn te verklaren door de achtergrondkennis (zowel cartografisch als domeinspecifiek) van de experten. Het is van cruciaal belang om dit verschil te begrijpen, zodat er richtlijnen kunnen opgesteld worden voor de cartografen (experten) die rekening houden met de (cognitieve) beperkingen van de eindgebruikers van de kaarten (vaak niet-experten). Anderzijds kunnen deze richtlijnen worden geïmplementeerd in de tools waarmee niet-experten tegenwoordig cartografische producten kunnen aanmaken op het internet, zodat deze eindproducten nog steeds op een effectieve manier informatie kunnen overbrengen naar de eindgebruikers.

Referenties


Curriculum Vitae

Kristien Ooms was born in Hasselt (Belgium) on the 2\textsuperscript{nd} of August in 1984. In 2002 she graduated at the Sint-Ursula Instituut in Herk-de-Stad and started her academic education at the Department of Geography (Ghent University) in that same year. She obtained her Master’s degree (\textit{magna cum laude}) in Geomatics and Surveying in 2006 and a Master’s degree (\textit{magna cum laude}) in Applied Informatics in 2007.

From June till December 2007, Kristien worked at the department of Geography on a project of the Belgium national mapping agency (NGI). In January 2008, she started working as an assistant at the department. In this function she was involved in several courses in the field of cartography and geographic information systems, such as General and Applied Cartography, Computer Cartography, Map Projections and Coordinate Systems, Geomatics Programming, etc.

Kristien contributed to several international conferences, where her presentations were evaluated very positively by the scientific cartographic community. As a consequence, she was nominated to be a vice-chair of the Commission on Use and User Issues of the ICA (International Cartographic Association), which was approved in 2011. Kristien is the author of a number of papers that are published in leading international journals in the field of cartography and GIS.
In this dissertation, Kristien Ooms investigate how novice and expert map users read, interpret, store, and use the visual information presented on screen maps. A combination of techniques is applied during a number of subsequent user studies, which focussed on different aspects on the users’ cognitive processes. The obtained eye tracking data, reaction time measurements, thinking aloud protocols, sketch maps, and questionnaires are analysed quantitatively and qualitatively. These insights are used to formulate recommendations for an effective map design.