Assessing Visual Variables of Cartographic Text Design

Rasha Deeb

geography department
Ghent University

UNIVERSITEIT GENT

Assessing visual variables of cartographic text design

Rasha DEEB

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

Cover design: Wim VAN ROY, Ninar AL-BARRI, Rasha DEEB

Faculty of Science
Department of Geography

Assessing visual variables of cartographic text design

Dissertation submitted in accordance with the requirement for the degree of doctor of science: Geography

Het beoordelen van visuele variabelen in cartografisch tekstontwerp

Proefschrift aangeboden tot het behalen van de graad van Doctor in de Wetenschappen: Geografie

Rasha Deeb

Supervisors

Prof. dr. Philippe De Maeyer
Department of Geography, Ghent University

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

Dr. Kristien Ooms
Department of Geography, Ghent University

Members of the Reading Committee

Prof. dr.ir Temenoujka Bandrova
Laboratory on Cartography ,University of Architecture, Civil Engineering and Geodesy, Sofia

Dr. *Corné van Elzakker*
Department of Geo-Information Processing, International Institute for Geo-Information Science and Earth Observation, University of Twente

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

Members of the Examination Committee

Prof. dr. Ben Derudder (Chairman)
Department of Geography, Ghent University

Prof. dr. Necla Ulugtekin
Department of Geomatics Engineering, Istanbul technical University

Table of Contents

Table of Contents V
List of Figures XI
List of Tables XV
List of Abbreviations XVII
List of Accordance XIX
Acknowledgement XXI
1 1
Introduction 1
1.1 Overview 1
1.1.1 Visual variable definition, development, and limitations 2
1.1.2 Principles of visual variables' use 7
1.1.3 Cartographic text design in focus 9
1.1.4 Measuring a good label design 12
1.2 Rationale and synopsis 14
1.2.1 Research objectives and questions 14
1.2.2 Outlines of the dissertation 17
1.2.3 Out of scope 21
1.3 Reference 22
2 27
Cartographic Text Preference, Gender and Expertise Variation 27
Preface 27
2.1 Background and Objectives 28
2.2 The Methodology 31
2.2.1 Participants 31
2.2.2 Task, stimuli and data 31
2.3 Results and Discussion 33
2.3.1 Measurement of preference and typographic quality for all users 33
2.3.1.1 Size 33
2.3.1.2 Shape 34
2.3.1.3 Orientation 35
2.3.1.4 Texture 36
2.3.2 Experience and gender influence 37
2.3.2.1 Experts versus novices 37
2.3.2.2 Females versus males 37

2.4 Conclusions and Future perspectives 39
2.5 References 41
3 43
Toward a Deeper Understanding of Cartographic Text Visualisation: Assessment of User Preferences and Colour Influence 43
Preface 43
3.1 Introduction 44
3.2 Method 46
3.2.1 Participants 46
3.2.2 Stimuli 46
3.2.3 Map design 47
3.2.4 Task 48
3.3 Results 48
3.3.1 Labels associated with areal data 48
3.3.1.1 Size of the labels 48
3.3.1.2 Shape of the labels 51
3.3.1.3 Texture of the labels 52
3.3.1.4 Orientation of the label. 53
3.3.2 Labels associated with point data 54
3.3.2.1 Size and shape variations 55
3.3.2.2 Texture of labels 55
3.3.2.3 Combining the variations of labels 56
3.4 Discussion 56
3.5 Conclusions and Future work 57
3.6 References 58
4 61
Evaluating the Efficiency of Typographic Design: Gender and Expertise Variation 61
Preface 61
4.1 Introduction 62
4.1.1 Label Design 62
4.1.2 Map Usability Testing 63
4.2 Study Design 64
4.2.1 Participants 64
4.2.2 Task and Stimuli 64
4.3 Results and Discussion 66
4.3.1 Data 66

4.3.2 Size of Labels 66
4.3.3 Shape of Labels 69
4.3.4 Orientation of Labels 71
4.3.5 Texture of Labels 72
4.4 Conclusions and Perspectives 75
4.5 References 76
5 79
Cartographic Text, the Preference of Label Design, and Cross-Cultural Comparisons 79
Preface 79
5.1 Introduction 80
5.2 Study design 80
5.2.1 Apparatus 80
5.2.2 Participants 81
5.2.3 Maps and typographic design 81
5.2.4 Stimuli, task and procedure 82
5.2.5 Data and recordings 82
5.3 Results 82
5.3.1 Size 82
5.3.1.1 Latin versus Cyrillic 82
5.3.1.2 Latin, Arabic, and Chinese 83
5.3.2 Shape 84
5.3.2.1 Latin versus Cyrillic 84
5.3.2.2 Latin, Arabic, and Chinese 84
5.3.3 Orientation 86
5.3.3.1 Latin versus Cyrillic 86
5.3.3.2 Latin, Arabic, and Chinese 87
5.3.4 Texture 89
5.3.4.1 Latin versus Cyrillic 89
5.3.4.2 Latin, Arabic, and Chinese 90
5.4 Discussion 90
5.5 Conclusion and Future work 92
5.6 References 92
6 93
Cartographic Text, The Efficiency of Label Design, and Cross-Cultural Comparisons 93
Preface 93
6.1 Introduction 94

6.2 Characteristics of different lettering systems ... 94
6.3 Study design ... 97
6.3.1 Participants ... 97
6.3.2 Stimuli ... 98
6.3.3 Task ... 99
6.3.4 Procedures and apparatus ... 99
6.3.5 Recordings ... 100
6.4 Results ... 100
6.4.1 Size ... 101
6.4.1.1 Within lettering system size analysis ... 101
6.4.1.2 Between lettering systems size analysis ... 102
6.4.2 Shape ... 102
6.4.2.1 Within lettering system shape analysis ... 103
6.4.2.2 Between lettering systems shape analysis ... 104
6.4.3 Orientation ... 104
6.4.3.1 Overall label orientation ... 104
6.4.3.1.1 Within lettering system overall orientation analysis ... 104
6.4.3.1.2 Between lettering systems overall orientation analysis ... 105
6.4.3.2 Characters' orientation: italic ... 105
6.4.3.2.1 Within lettering system characters' orientation analysis ... 106
6.4.3.2.2 Between lettering systems characters' orientation analysis ... 106
6.4.4 Texture ... 107
6.4.4.1 Within lettering system texture analysis ... 108
6.4.4.2 Between lettering systems texture analysis ... 109
6.5 Discussion ... 110
6.5.1 Label design efficiency ... 110
6.5.2 Cross cultural comparison ... 111
6.5.3 Limitation and Generalizability ... 112
6.6 Conclusion ... 112
6.7 Appendix-1 ... 113
6.8 References ... 115
7 ... 117
Assessment of Map Users' Search Strategies Using Multi-Package Software for Scanpath Analysis ... 117
Preface ... 117
7.1 Introduction ... 118

7.2 Study design ... 119
7.2.1 Apparatus ... 119
7.2.2 Participants ... 119
7.2.3 Task and stimuli ... 120
7.2.4 Procedures ... 120
7.3 Results & Discussion ... 122
7.3.1 Data processing, simplification, and scoring ... 123
7.3.2 Output and visualization ... 124
7.3.2.1 Expertise variations ... 126
7.3.2.2 Gender variations ... 126
7.3.3 Clusters' description ... 129
7.4 Conclusion ... 130
7.5 References ... 130
8 ... 133
Background and Foreground Interaction, and the Influence of Complementary Colours on Search Tasks ... 133
Preface ... 133
8.1 Introduction ... 134
8.2 Study design ... 135
8.2.1 Apparatus ... 135
8.2.2 Participants ... 135
8.2.3 Map design ... 136
8.2.4 Stimuli, task and procedure ... 137
8.2.5 Data and recordings ... 138
8.3 Results ... 139
8.3.1 Response time measurements ... 139
8.3.2 Fixation duration ... 140
8.3.3 Fixation count (fix/s) ... 140
8.3.4 Colour difference and luminance contrast ... 141
8.3.5 Post-test questionnaire ... 145
8.4 Discussion ... 145
8.5 Conclusion and future work ... 146
8.6 References ... 146
9 ... 149
General Discussion ... 149
9.1 Summary ... 150

9.2 Improving cartographic text design on screen maps 158
9.3 Limitation, challenges, critical reflections and avenues for future research 159
9.4 References 160
10 163
Conclusions 163
11 165
Summaries 165
11.1 Summary 166
11.2 Samenvatting 172
Questionnaires 179
Curriculium Vitae 183

List of Figures

Figure 1.1: The visual variables and their application to the cartographic text 6
Figure 1.2 : Map–use cube. Goals and activities of map use (patterned after MacEachren and Kraak, 1997) 8
Figure 1.3 : Outline of the dissertation 17
Figure 1.4: The relationships between chapters of the dissertation 20
Figure 2.1: Example of the ranking task for point data (variable: texture) 32
Figure 2.2: Example of the ranking task for areal data (variable: size) 32
Figure 2.3: Shape preference of the same texture 35
Figure 2.4: The preference of shape and size of different textures 36
Figure 2.5: Significant differences between experts and novices 38
Figure 2.6: Located significant differences for females and males... 39
Figure 3.1: A comparative example of test stimuli of the four background types, Areal labelling (left) and point labelling (right) 47
Figure 3.2: The mean users' preferences on both size groups (normal & bold) regarding the four dot sizes 49
Figure 3.3: The mean values of users' preferences on size (normal and bold) regarding the maps background colours 49
Figure 3.4: Trends of the frequency of users' preferences of labels' size regarding to background colour 50
Figure 3.5: Users' preference on shape categories regarding background colour 51
Figure 3.6: The mean values of users' preferences of texture variable.. 52
Figure 3.7: The influence of background colour onto the mean preference of texture variable 52
Figure 3.8: The mean user's preference on label orientation (straight versus italic) for the four backgrounds' colour tests 54
Figure 3.9: The influence of background colour onto the mean preference of label orientation (straight versus italic) 54
Figure 3.10: The influence of background colours onto the mean user's preference of labels' shape including case style variations 55
Figure 4.1: An example of two test screens, (a) point data labelling, (b) areal data labelling 65
Figure 4.2: An example of label target distribution on the topographic map 65
Figure 4.3: Box plots of the reaction time measurements (in ms) 68
Figure 4.4: Box plots of the reaction time measurements (in ms) 73
Figure 4.5: a comparision between 2 levels (left) and 3 levels (right) of hierarchy with Arial shape .. 74
Figure 5.1: Both bold and normal size comparison for Cyrillic and Latin lettering systems................ 83
Figure 5.2: Normal size comparison to bold sizes for Latin, Arabic, and Chinese lettering systems... 84
Figure 5.3: Normal size comparison to bold sizes for Latin, Arabic, and Chinese lettering system 85
Figure 5.4: The application of Arial shape on both Latin and Cyrillic lettering system...................... 86
Figure 5.5: The application of the overall label orientation on both Latin and Cyrillic lettering system 87
Figure 5.6: Users' preference of Label orientation for the entire lettering system 87
Figure 5.7: Normal and italic comparison for Latin, Arabic, and Chinese lettering systems 88
Figure 5.8: Users' preference of Labels' texture (sans serif versus serif) for both Latin and Cyrillic . 89
Figure 5.9: Users' preference of Labels' texture (normal versus bold) for both Latin and Cyrillic 90

Figure 5.10: Users' preference of Labels' texture (normal versus bold) for Latin, Arabic, and Chinese 90
Figure 6.1: The contextual shape of Arabic letters (isolated, initial, middle, and final) 95
Figure 6.2: The differences made on the word caused by ligature (left) and kashida (right) 96
Figure 6.3: Example of the parallel map design for the two map (left versus right) types and three lettering systems (top: Latin; middle: Arabic; bottom: Chinese) 98
Figure 6.4: Example of the test stimuli with the Arabic lettering system 100
Figure 6.5: Size comparison of Chinese typography; the first row is normal size, and the second row is bold 101
Figure 6.6: Mean reaction time measurements for the three lettering systems when different label sizes are used (left: normal labels; right: bold labels); the error bars represent the standard deviation 102
Figure 6.7: An example of both shape A (left) and shape B (right) for the three lettering systems 103
Figure 6.8: Mean reaction time measurements for users of different lettering systems and for different shapes; the error bars represent their standard deviation. 104
Figure 6.9: Small sketches of maps representing the overall label orientation 105
Figure 6.10: Mean reaction time measurements for users of different lettering systems and for different overall orientations; the error bars represent the standard deviation 105
Figure 6.11: Mean reaction time measurements for italic orientation for users of different lettering systems over four consecutive sizes; the error bars represent their standard deviation 107
Figure 6.12: An example of two different textures: left, serif normal size hierarchy; right, bold size hierarchy (normal versus bold design, shape B) 107
Figure 6.13: Mean values of users' reaction time measurements for the hierarchy of the three lettering systems for both the normal and bold sans serif-shape B-sizes, (left) normal and (right) bold; the error bars represent their standard deviation 108
Figure 6.14: Mean values of users' reaction time measurements for the hierarchy of the three lettering systems for both the normal and bold serif-shape B-sizes, (left) normal and (right) bold; the error bars represent their standard deviation 108
Figure 7.1: The targets distribution over the stimuli dimension. Targets are surrounded with white rectangles 119
Figure 7.2: Data treatment procedures, using the combination of different software packages 120
Figure 7.3: Areas of Interests' design according to the spatial structure of the stimuli 121
Figure 7.4: An example of four scanpaths for the same stimulus 122
Figure 7.5: An example of scanpath comparison between two groups, (left: experts versus novices, right: females versus males) 122
Figure 7.6: Left, visual representation of tree branches showing the degrees of similarities in search systems for expert group generated in EyePatterns, the first stimulus is represented here; right, a redrawn hierarchal pattern by the authors based on the relationships in the original tree 125
Figure 7.7: Users' performance on 6 stimuli (1-6) considering the expertise and Gender variations, experts (grey) versus novices (white) and females (red) versus males (blue) 127
Figure 7.8: Users' performance on 6 stimuli (7-12) considering the expertise and Gender variations, experts (grey) versus novices (white) and females (red) versus males (blue) 128
Figure 7.9: Scanpath patterns in OGAMA, male cluster in stimuli 9 129
Figure 7.10: Scanpaths pattern in EyePatterns. Users' performance on stimulus 9, gender variation 130
Figure 8.1: The twelve tested colour with their common names 136
Figure 8.2: Target label distribution of the map face 137
Figure 8.3: the stimuli design with an example of map order 138

Figure 8.4: Scatterplots showing the influence of the colour difference on the users' performance and the related regression model 142
Figure 8.5: Scatterplots showing the influence of the colour difference in relation to the average fixation count per second (up); colour difference between the foreground and background for both coloured and black labels (down) 143
Figure 8.6: Scatterplots showing the influence of the luminance difference in relation to the target fixation duration (up); the difference between the target luminance and background luminance for the 24 map designs (down) 144

List of Tables

Table 1.1: The effectiveness of applying visual variables to cartographic text in relation to the associated visualized data. 9
Table 2.1: Levels of organization of the visual variables. After Bertin (1967, fig. 2, p.69) 28
Table 2.2: Bertin's variables on text 29
Table 2.3: Size overall preference 34
Table 2.4: Boldness effect on text preference 34
Table 2.5: The third level of complexity preference (with 2 and 3 levels of hierarchy) 34
Table 2.6: The influence of shape on text preference 35
Table 2.7: Orientation influence on text preference 36
Table 2.8: Complexity effect on orientation of text preference 36
Table 2.9: The detected level of preference variation between experts and novices 37
Table 2.10: The detected significance levels for gender preference variations 38
Table 3.1: HSV parameters of the map designs 48
Table 3.2: Users' preference on label orientation related to background colour 53
Table 3.3: Percentage of users' preference according to case variation and font type 55
Table 4.1: Mean reaction times (in ms) for males and females, related to label sizes 67
Table 4.2: Mean reaction times (in ms) for novices and experts, related to label sizes 67
Table 4.3: Mean reaction times (in ms) for males and females, related to the labels' shape 69
Table 4.4: Mean reaction times (in ms) for novices and experts, related to the labels' shape 69
Table 4.5: Mean reaction times (in ms) for males and females, ... 70
Table 4.6: Mean reaction times (in ms) for novices and experts, ... 70
Table 4.7: Mean reaction times (in ms) for males and females, related to the labels' orientation 71
Table 4.8: Mean reaction times (in ms) for novices and experts s, related to the labels' orientation ... 71
Table 4.9: Mean reaction times (in ms) for males and females, related to the use of italic 72
Table 4.10: Mean reaction times (in ms) for novices and experts, related to the use of italic 72
Table 4.11: Mean reaction times (in ms) for males and females, ... 74
Table 4.12: Mean reaction times (in ms) for novices and experts, ... 74
Table 5.1: Bertin's variables on text 81
Table 5.2: typographic size comparison for both Latin and Cyrillic (vertical versus horizontal). 83
Table 7.1: Users' fixation average per stimulus and fixation sum considering their characteristics .. 121
Table 7.2: Two scanpath analysis, using rewarding and penalizing 123
Table 7.3: The scanpath similarity between experts and novices in the first stimulus calculated by OGAMA, experts are placed vertically and novices horizontally 124
Table 7.4: The scanpath similarities between experts (A) and novices (B) in the first stimulus presented by OGAMA 125
Table 8.1: The used colour in the stimuli presented in HSV and RGB when the maps were designed and the measured CIE and XYZ systems when the maps were displayed. 137
Table 8.2: A comparison of the users' responses (s) between black and coloured labels (M= Mean, SD= Standard Deviation). 139
Table 8.3: The statistical model of the MANOVA test regarding the map design, users' expertise and users' gender for reaction time measurements(s), fixation duration (s) and fixation count (Fix/s). ... 140
Table 9.1 : The head line of the dissertation's component 151

List of Abbreviations

2D	Two Dimensions
2.5D	Two and half Dimensions
3D	3 Dimension
HSV	Hue, saturation, and Value colour system
RGB	Red, Green, and Blue colour system
CYMK	Cyan, yellow, Magenta, and Black colour system
CIE	Colour spaces issued by international commission on illumination (Commission International de l'Éclairage)
ISO	the International Organization for Standardization
RQ	Research Question
TNR	Times New Roman
ANOVA	ANalysis Of VAriance
BM	Blank Map
GS	Grey Scale
HC	Hot Colour
CC	Cold Colour
F	Degree of freedom
P	Significance difference
RT	Reaction Time
KB	Kilo Bite
M	Mean
SD	Standard Deviation
MS	MilliSecond
Hz	Hertz
F_L	Degree of freedom concerning Latin group
P_L	Significance difference concerning Latin group
F_A	Degree of freedom concerning Arabic group
P_A	Significance difference concerning Arabic group
F_C	Degree of freedom concerning Chinese group
P_C	Significance difference concerning Chinese group
M_L	Mean value concerning Latin group
M_A	Mean value concerning Arabic group
M_C	Mean value concerning Chinese group
AOI	Area Of Interest
ICC profile	International Color Consortium's color attributes of a particular device
CIE Lab	A *Lab* color space is a color-opponent space with dimension *L* for lightness and *a* and *b* for the color-opponent dimensions
CIE XYZ	Tristimulus colour space based on Y as colour Luminance
MANOVA	Multi ANalysis Of VAriance

List of Accordance

Chapter Number	**Chapter Accordance**	**Chapter Reference Modified from:**
Chapter 2	**Deeb *et al.*, 2012**	Deeb, R., Ooms, K. & De Maeyer, P., 2012. Typography in the Eyes of Bertin, Gender and Expertise Variation. The Cartographic Journal, 49(2), pp. 167-185.
Chapter 3	**Deeb *et al.*, 2013b**	Deeb, R., Ooms, K., Van Etvelde, V. & De Maeyer, P., 2013b. Toward a deeper understanding of cartographic text visualisation: assessment of user preferences and colour influence. The Cartographic Journal.Accepted paper.
Chapter 4	**Deeb *et al.*, 2013a**	Deeb, R., Ooms, K., Vanopbroeke, V. & De Maeyer, P., 2013a. Evaluating the Efficiency of Typographic Design: Gender and Expertise Variation. The Cartographic Journal. 51(1), pp. 75-86.
Chapter 5	**Deeb *et al.*, 2014b**	Deeb, R., Bonchev, S., Bandrova. T.,& De Maeyer, P., 2014a. How to design users' friendly multilingual maps? Cyrillic and Latin Labels. 5th International Conference on cartography and GIS. Riviera.
Chapter 5	**Deeb *et al.*, 2015c**	Deeb, R., Kurban, A., & De Maeyer, P. 2015c. From Monolingual to Multilingual Mapping: What is users' preference?. 27th International Cartographic Conference, Rio de Janeiro. Accepted – ICC 2015.
Chapter 6	**Deeb *et al.*, 2015a**	Deeb, R., Ooms, K, Kurban, A., Ablikim, A., & De Maeyer, P. 2015a. Cartographic text, the efficiency of label design and cross culture comparisons. Information Visualization. Submitted 2015.
Chapter 7	**Deeb *et al.*, 2015b**	Deeb, R., Ooms, K., Popelka, S., Van Etvelde, V. & De Maeyer, P. 2015b. Assessment of map users' search strategies using multi-packages software for scanpaths analysis. 27th International Cartographic Conference, Rio de Janeiro. Accepted – ICC 2015.
Chapter 8	**Deeb *et al.*, 2014b**	Deeb, R., Ooms, K., Van Etvelde, V. & De Maeyer, P., 2014b. Background and foreground interaction: complementary colours' influence on search task. Color Research and Applications. Accepted 28 August 2014.

Acknowledgement

Since I was a kid, I had the faith that there will be a day when I will travel abroad and pursue a PhD research. I followed my dream, and here I am, doing the final steps to obtain this degree. The road to reach this dream was not paved; it was rather full of obstacles and difficulties. At this stage of the road, I look back and I remember several people, who showed me that there is always a candle to light the darkest roads.

I would like to present my gratitude and appreciations to my supervisor Philippe De Maeyer. I still remember that day when I received an e-mail from him in which he approved to supervise me as PhD student in Ghent University. I believe that was a turning point in my life. Words cannot show my thanks for giving me the opportunity to start PhD degree, guiding me along the way, having faith in me, for his wisdom and tireless assistance. In addition to cartography, without him knowing, Philippe taught me lessons in life. Philippe....you were not only a supervisor but also you were a father to me.

Many thanks also go to my co-promoter Veerle Van Eetvelde for her help, support and advices that she provided me with to enhance my work and to illuminate on some aspects of my work

I shall also thank my entirely patient co-promoter, Kristien Ooms, who tirelessly edited several drafts of my papers, and I want to thank her for all her tedious work and extensive guidance.

The Godmother of all ‘Helga’, the department secretary. For me, she was not only a secretary who is taking care of the department inputs and outputs, but she was the wise whisper who helped me to integrate well in the university and in Ghent. I will take the opportunity to thank the supporting staff, Nathalie, Karine, Sabine, Wim, Paul, Iris.

Although this dissertation is named by one author, but it would not be possible to accomplish without the help and support from my office colleagues Ann and Soetkin (Tata and Kotken ☺) who shared may laughs and tears. Thank you for being my friends. Without you both, I cannot imagine how desperate it will be to come to the office every day. I will also take the chance to thank my friends; Katrien, Timo, Niels, Seann, Mathias and Coen. Many thanks for all my colleagues in the department for their support and for taking part in this research’s experiments.

From very far away and over the seas, my family have been sending the spiritual courage for me to continue the road. I know that you are proud of me but you must know that what I have accomplished so far is meaningless without you. Finally my eternal thanks go to my husband Bashar, thank you for being my home…..Oh ..Oh …oops. I was about to forget thank you Ninar...life would have been so miserable without your smile.

Rasha

18.2.2015

1

Introduction

1.1 Overview

Geographic Information System (GIS) and mapping software packages have significantly facilitated production of both paper and screen maps. Additionally, data collection techniques and methods have carried maps into their golden era (Kraak and Ormeling, 2010). However, with web mapping, design issues occur regarding communication with users (MacEachren, 1995).

Cartographers employ their knowledge and experience to translate the physical world into visual representations using the cartographic language. The map user decodes the cartographic language according to his cartographic knowledge and recreates and interoperates the depicted reality. Many factors can influence this process, such as knowledge, experience, psychological processes, abilities, aims, needs, interests, tasks, skills and other external conditions. Therefore, the reality created on the users' side will not match the reality created initially by the cartographer. This fact causes errors in the communication process (Kolácný, 1969).

It is likely that a 'poorly' designed map leads to incorrect communication on the users' side, and the 'well'-designed map results in better communication with users. However, the attributes of 'bad' and 'good' map design are not only determined by cartographic theoreticians and designers but also by map users (Board and Taylor, 1977; Kraak and Ormeling, 2010).

The mechanism of how map users interpret maps has been discussed by many cartographers and the understanding of the sign system, its development, and its use has been the goal of many research projects in thematic cartography and geographic visualization (Robinson, 1952; Bartz Petchenik, 1974; Board and Taylor, 1977; MacEachren and Taylor, 1994; MacEachren, 1995; Hinks, 2009; and others).

Bertin's[1] contribution to the graphic world is one of the earliest (1967) and the most important; it is considered the basic root of sign systems and design studies and has been the main reference for cartographers since then. This research will focus on what Bertin presented to define and develop graphics and will consider his variables' applications on cartographic texts and their influence on users' preferences and efficiency.

Bertin stated in the beginning of his book, Semiology of Graphics (1967), 'Graphic representation constitutes one of the basic sign-systems conceived by the human mind for the purpose of storing, understanding, and communicating essential information. As a "language" for the eye, graphics benefits from the ubiquitous properties of visual perception'. This statement focuses on two basic elements:

- how the sign system is organized, perceived, interpreted, and comprehended by the human mind. The information's level of organization within our mind is linked to data type and classification, which can be either quantitative or qualitative. This issue will be discussed in section 1.1.2; and
- visual variables (at a fixed X,Y position, Bertin's variables are size, shape, colour-hue, value, orientation and texture; they will be discussed thoroughly in the following section) and their properties' influence on the visual perception of associativity, selectivity, order and quantity. Properties of visual variables will be explained in 1.1.2.

1.1.1 Visual variable definition, development, and limitations

The linguistic definition of visual variables is 'changes that are seen.' The Oxford dictionary defines visual as 'of beams coming, proceeding, or directed from the eye or sight'; and variable as 'what is liable or apt to vary or change; (readily) susceptible or capable of variation; mutable, changeable, fluctuating, uncertain.' The Cambridge dictionary explained the term by saying that visual is 'what relates to seeing' and variable is 'either a number, amount, or situation that can change.'[2]

When Bertin (1967) first issued the term visual variables he also named them retinal variables. He attributed the visually variable mark on two planar dimensions by saying, 'fixed at a given point, the mark, provided it has a certain dimension, can be drawn in different mode by varying in size,

[1] In this dissertation, the translated version of Bertin's book 'sémiologie graphique: Les diagrammes-Les réseaux-Les cartes' was used. The translation was done by Esri press, 2011.
[2] To look up the term's linguistic meaning, both Oxford and Cambridge dictionaries were used on 14th of October 2014.

value, texture, colour, orientation, and shape.' Later, Slocum *et al*. (2005) used the term visual variable to describe the various perceived differences in map symbols that are used to present geographic phenomena.

(Bertin, 1967) set some limitations for his representations using visual variables in the graphic-sign system:

- 'representable or printable;
- on a flat sheet of white paper;
- standard size, visible at glance;
- at a distance of vision corresponding to reading from a book or atlas;
- under normal and constant light;
- utilizing readily available graphic means.'

However, he excluded:

- 'Variation of distance and illumination, i.e., all display materials that are different from a sheet of paper;
- Actual relief (thicknesses, anaglyphs, stereoscopics), i.e., 3D representation;
- Actual movement (flicker of movement of the image, animated drawing, films).'

Some more limitations of Bertin's variables arose with the developments of new visualization techniques, tools, and display media (from large screens to mobile maps), especially with the noticeable transition from paper to screen maps. However, Bertin's theory has proven to be very adaptable and open to innovation (Koch, 2001). As an example, the X, Y positioning variable in 2D visualization scaled up to X, Y, Z positioning in 3D visualization. Meanwhile, MacEachren (1994) linked the delineations of Bertin's variables to the methodology, which creates the connection between the sign and its meaning and the medium in which the map is displayed. Although Bertin's theory proved to be open to innovation (Koch, 2001), many dynamic visual variables were added to the designs' elements of cartography: moment (display time), duration, frequency, order, rate of change and synchronisation (DiBiase *et al.,* 1992; MacEachren, 1994; Köbben and Yamman, 1996; Kraak and Ormeling, 2010). In addition to the mentioned temporal elements, some modern cartographers promote the use of sound as an extra aural variable that adds value to the visual variables by improving their perception (Koch, 2001; Kornfeld *et al.,* 2011) and further define a sound's dimensions as pitch, timber, volume, speed, and location (Dransch 1995; Koch,2001).

The evolution of digital geospatial data handling gained momentum in the 1980s. Since then, more and more software packages were developed for cartography, for the purpose of both the production and the use of maps. Thus, users had two type of maps based on the presentation medium: paper maps and digital maps. Paper maps play the traditional role of providing an overview of the geospatial relationships and patterns, while digital maps can also function as interfaces or indexes for extra information (Kraak, 2001). This dissertation will consider how Bertin's variables, which were initially proposed for the semiotics of paper maps, can be applied to improve digital map design.

Until now, the rather dogmatic nature of Bertin's theory has been discussed. Some cartographers blindly used it as a standard reference for their designs. Others encountered the implementations of modern cartography (digital, dynamic, and interactive cartography) when applying the theory. When some cartographers took a closer look at the perceptual criteria of each of Bertin's variables in the light of new psychological theories or cartographic development, they introduced some variables as extra tools to enhance symbol design, providing rules of legibility (graphic density, angular separation, retinal separation). Morison (1974) added arrangement and a third dimension for colour (saturation was added to hue and value). To adjust Bertin's 2D variables, Slocum *et al.* (2005) proposed the use of spacing, size, perspective height, orientation, shape, arrangement and finally hue, lightness and saturation for 2.5D and 3D representations.

The characteristics of Bertin's visual variables and how they can be perceived and comprehended are demonstrated in Figure 1.1 and presented below:

I. **Size**:

Differences in symbol size can refer to degrees of symbols' height, diameter, and diagonal, or variations in symbols' areas. The size difference can be either over the whole symbol or over the individual marks which form the symbol (this will be considered textural variation, as will be explained later, or texture size in Figure 1.2). When designing topographic and thematic maps, size plays a major role in expressing the order relation of the objects' class, for example the hierarchy of routes or the hierarchy of cities (Denegre, 2005).

Size is correlated to symbol dimensions in relation to the map scale. However, this concern is a growing issue when designing both static and interactive digital maps where size is also linked to map resolution, as elements' size is defined by their number of pixels. See the first row in Figure 1.1.

II. **Shape**:

Within a fixed size, an infinite number of forms can be presented; the approach can be systematic or non-systematic, geometric or non-geometric. Shapes are determined by the contour or boundary between the object and the non-object (MacEachren, 1994), represented in the second row of Figure 1.1.

Gestalt psychology will help to define forms by presenting several gestalt principles. Some of the gestalt laws, which influenced cartographic designs, are explained below (Moore and Fitz, 1993; Hinks, 2009; Arntson, 2011):

- pragnanz(conciseness, good form): ambiguous or complex images or patterns are simplified and completed by the viewer to the nearest and complete shape;
- proximity: objects that are close to each other are perceived to be more related than objects that are spaced farther apart. Thus, viewers tend to interpret objects near each other as a group or as a perceptual unit;
- similarity: objects that are similar are perceived to be more related than dissimilar objects, and similar objects are likely to be grouped together;
- symmetry and balance: viewers perceive information (stimuli) to be incomplete if they are not balanced or symmetrical;
- closure: when looking at a complex arrangement of individual elements, observers tend to first look for a single, recognizable pattern, i.e., full and complete visual data are easier to interpret than incomplete data or patterns, which can cause disturbance to the viewer;

- continuity: observers are keen to follow a certain path when looking to neighbouring objects. This connection can easily be identified as their eyes will follow an obvious path (aligned on a straight or well-defined curved line) or the flow of information;
- common fate: observers tend to perceive objects moving in the same direction, at the same pace, at the same time as being more related than stationary objects or objects that move in different directions;
- familiarity: objects are more likely to form groups if the groups appear familiar or meaningful;
- figure-ground segregation: elements are perceived as either figures (distinct elements of focus) or ground (the background or landscape on which the figures rest);
- focal point: this is a point of interest, something emphasized or different. A focal point will capture the viewer's attention; and
- harmony (unity): objects that share uniform (visual connection) or visual characteristics are perceived as being more related than elements with disparate visual characteristics.

III. **Colour:**

The perception of wavelength in the visible light. Colour can be defined in three dimensions:

- **Colour hue**: the dominant wavelength which can be perceived by the eye, illustrated in the third row of Figure 1.1;
- **Colour value** (lightness): at a fixed hue, value refers to how light or dark a colour can be. Thus, a value of zero results in a black appearance and a value of 100 (maximum value) results in a white appearance. Therefore, a scale for each colour hue is formed from 0 to 100, where hue ranges from black to white values, as seen in the fourth row of Figure 1.1;
- **Colour saturation**: the intensity of colour. The saturated tone is not of a constant value but varies in value according to the colour (Bertin, 1967). Saturation is also called purity, intensity, or chroma (Tyner, 2010). Saturation is illustrated in the fifth row of Figure 1.1.
- Bertin clearly considered colour hue and value to be visual variables. Although he discussed saturation, he did not consider it a visual variable because its complications in graphic representation are numerous. To provide accurate colour definitions, the age of digital cartography (and even before) enforced the use of colour in three dimensions. Moreover, digital cartography has introduced different colour coordinate systems: HSV, RGB, CMYK, and CIE.

Kraak and Ormeling (2010) defined four aspects of colour perception, which can be used to improve colour perception in relation to other visual variables. These aspects shall be considered when planning coloured maps (Tyner, 2010):

- **The psychological aspects**: this refers to colour preference and its qualities, relation to human behaviour, cultural background, and association between colour and mood. For example, it is known that warm colours can be motivating and cold colours can be inhibiting;
- **The physiological aspects**: some physiological issues of colour perception shall be considered, such as the difficulty of perceiving colours on small areas; the ease of perceiving contrasts between particular colours (Kraak and Ormeling, 2010); differences in humans' acuity for different hues (MacEachren, 1994) and the influence of gender

(differences of visual acuity between females and males), age (visual acuity drops with age), and colour blindness fall in this category;

- **Connotative aspect**: colours are associated with different connotations (functional, emotional, and social); for example, red is associated with love, danger, and glory; black is associated with grief, sophistication, and fear; white is associated with happiness, sincerity, and purity; and so on.
- **Conventional aspects**: in many fields of science there is a standardised use of colour. For map design and production, some geographic phenomena are associated with standard colours, such as blue for water bodies and green for forestry.

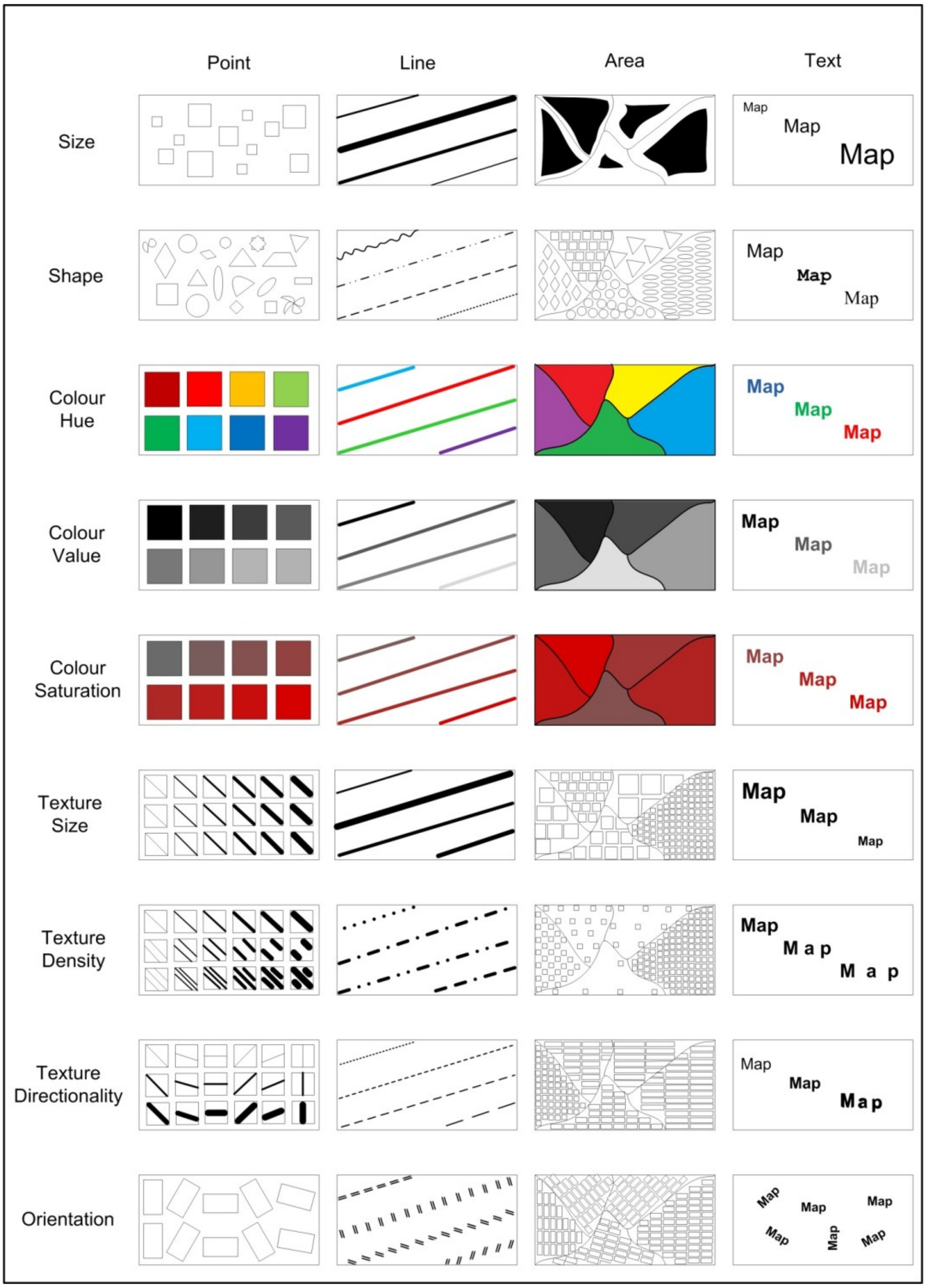


Figure 1.1: The visual variables and their application to the cartographic text

IV. **Orientation:**

A variable variable that describes the angle differences between symbols. Additionally, orientation functionally regulates the perception of direction, trend and movement. ‘If the direction of symbols has no significance to the distribution and is not a part of an areal pattern, then the symbol should be oriented or aligned with a common reference line, such as the map margins; otherwise, the reader assumes some significance for orientation’ (Tyner, 2010). This variable is illustrated in the last row of Figure 1.1.

V. **Texture (pattern, grain and arrangement**):

The term has been extensively discussed by cartographers and designers to produce a consistent definition. It describes the symbol when it could vary by more than one of the previous variables in a distinguishable manner: the way in which the marks are arranged and presented. Uttal (1988) emphasized the effect of marks’ organization on symbols’ perception as he demonstrated how visual perception is powerfully driven by the global organization of marks. Closure, harmony, and proximity of the gestalt laws explain how texture is perceived by viewers.

Caivano (1990) presented texture as a semiotic system. He did not simply treat texture as the spacing of pattern elements as Morison (1984) did, or as equivalent to Bertin’s grain, which can vary by the differences in fineness and coarseness of pattern elements, but as tripartite variable analogues to colour. Although Caivano’s system appears limited to area symbols, the distinctions he made are worth considering and can be applied to both point and line symbols. Three dimensions of texture are proposed: directionality (the ratio of length to width of texture units), size (of the texture units), and density (ratio of texture units to the background). Texture dimensions are illustrated in the sixth, seventh, and eighth rows of Figure 1.1. Caivano (1990), on the one hand, described the formation of simple textures by the repetition and juxtaposition of a minimal unity called unity of texture. It is composed of the texturing elements and their respective intervals. On the other hand, MacEachren (1994) proposed that texture be considered a higher-level visual variable consisting of units including shape, size, texture (in Bertin, ‘sense of grain’) and arrangement. Morrison (1984) proposed the term ‘arrangement’ to differentiate between regular and irregular texture. It is noticeable that MacEachren made a distinction between texture, pattern, grain, and arrangement. This distinction is a controversial issue in the perception of texture. In the present dissertation, the term ‘texture’ will be used for describing the visual variable that Bertin described. Additionally, the development of the term will be discussed.

1.1.2 Principles of visual variables’ use

The goal of visualization is to transform data into a perceptually efficient visual format (Ware, 2004). Map data are usually treated accurately to present a good communication product. To produce maps, both data visualization and communication are involved, as MacEachren and Kraak (1997) presented (see Figure 1.2). As they demonstrated, the map use can emphasize any of the use activities: exploration, analysis, synthesis, and presentation. How cartographers can communicate map data with the audience has been discussed since the 1960s, with many models of cartographic communication appearing to explain the process and how to evolve better perception (Kolácný, 1969; Board, 1972; Robinson and Petchenik, 1975; Board and Taylor, 1977; Shannon, 2001; and others). Providing the highest level of communication is one of the main targets of cartographers (MacEachren, 1995); research, visualization tools, techniques, strategies and applications are employed to meet this target.

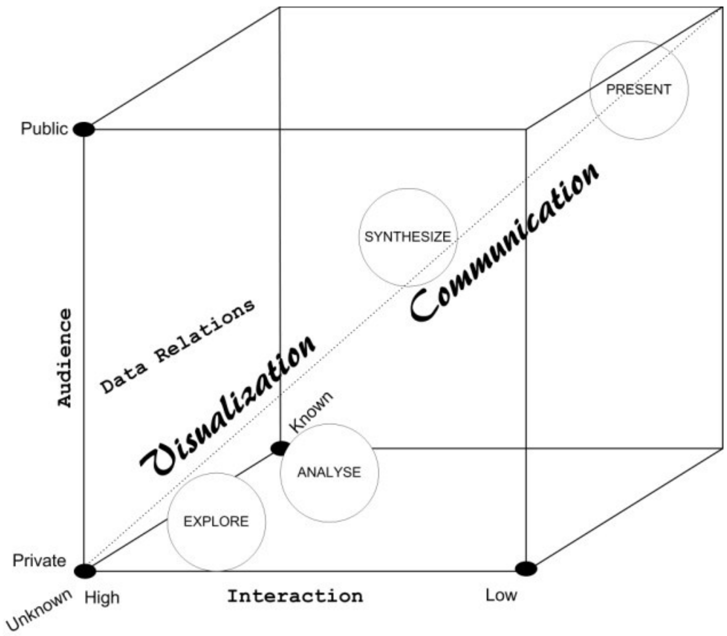


Figure 1.2 : Map–use cube. Goals and activities of map use (patterned after MacEachren and Kraak, 1997)

Previously, cartographers expected map users to more or less accept the cartographers' conditions and their designed maps. Users would learn how to work with any map the cartographers made (Robinson and Petchenik, 1975). However, the design of a map has a strong impact on how this map's information is perceived and interpreted, and thus has a great impact on the quality of the communication process (Robinson, 1952). Consequently, visual variables are used to visualize data and to help answer specific questions in relation to the geospatial data.

The level of organization and data type determines the use of visual variables. Notwithstanding that information's level of organization (Bertin, 1967), level of measurement (Slocum *et al.*, 2005), data measurements (Kraak and Ormeling, 2010), and level of classification (Tyner, 2010) are different titles that refer to the different types of geographic data, which can be any of the following:

- nominal data (qualitative);
- ordinal data (quantitative, and /or qualitative);
- interval data (quantitative);
- ratio data (quantitative).

Table 1.1demonstrates how the visual variables can represent effectively different type of data.

In regards the visualized data types, some visual variables are expected to be more efficient and effective than others for communicating data contents; for example, it is not advisable to use size to represent nominal data because the visual impact of size is linked to quantitative attributes (Poidevin, 1999). Additionally, the visual variables carry four perceptual properties (Bertin, 1967):

- **selectivity**: selective variables are used to answer the 'where' questions, giving the spatial identification of features and categories;

- **associatively**: this is helpful when we seek to equalize variation and to group correspondence within all categories of this variation;
- **order**: this property allow users to arrange and compare between two or more orders. It is not necessary to consult a map legend when variables are orders; and
- **quantity**: quantitative perception is useful to numerically define the ratio between two signs and to group homogeneous signs.

Table 1.1: The effectiveness of applying visual variables to cartographic text in relation to the associated visualized data (Derived from Bertin, 1967 and MacEachren 1995)

Attributed data	Qualitative Nominal	Ordinal	Quantitative Interval	Ratio
Size		+	+	+
Shape	+			
Colour hue	+			
Colour value		+	+	+
Colour saturation		+	+	+
Texture size		+	+	+
Texture density		+	+	+
Texture directionality	+			
Orientation	+			

1.1.3 Cartographic text design in focus

In one of the earliest cartographic works, which was presented by Withycombe in the Geographical journal in 1929, the essentials of typography on maps were rephrased as follows:

- **legibility**: the the letters must not only be legible when standing alone but also superimposed upon the details of maps;
- **suitability for reproduction by the photographic process which is to be used**: the names should not need retouching on the negative, should go down easily on the zinc plate, and should be free from any tendency to clog and thicken in rolling up.
- **good style and intrinsic decorative qualities**: style of the lettering on map should be as good as that exhibit by the fonts of type in use by book printers. As legibility is one of the characteristics of every really good alphabet, the first aim will be attained if really good style is achieved;
- **distinction and contrast**: certain classes of names should be clearly distinguished and different type of alphabets used and their gauge and spacing should achieve this.
- **harmony of effect:** the alphabets appearing on any one map should harmonize with each other, with the detail of map, and with the lettering of figures used in the margin for the title, imprint, reference, and so on.'

After this, little work was presented to guide cartographic text design. This concern became of high interest, especially with the development of map production software and techniques as well as the media on which maps were visualised.

Fairbairn (1993) explored the functionality of cartographic text and invited cartographers to consider cartographic text as the fourth symbol type (in addition to point, line, and polygon). He defined 15 functions of the cartographic text, which can attribute the mathematical component, the geographical component, and the complementary component of maps.

Text functions associated with the mathematical component of the map:

1- Geocoding

The geocoding is usually presented on the margins of the maps and usually associated with the grid (s). Conventional design recommends that the marginal texts have the same colour as the represented grid (Fairbairn, 1993).

2- Measurement

This function of cartographic text refers to the text associated with map scale. Additionally, description of the location of the inset maps falls in this category.

Text functions associated with the geographical component of the map:

1- Narrative Function (Labeling)

The narrative function is the most evident and used function of the cartographic text on mapface is to indicate the names of objects which are associated by point, linear, and polygon features.

2- Descriptive

To give identifications and additional properties to the portrayed features (e.g. seasonal rivers versus river). This function describes:

- **Warning**
  It is an additional form of the descriptive function and used to emphasis the dangerous nature of features. Such function can be presented by red colour to indicate warning in addition to the text itself.
- **Functional information**
  It is more specific manner than more descriptive text. It is presented by highlighting the items in the mapped area. This can be done in multiple ways such as variation in colour, presenting text in upper case, or using buffers around the text.
- Interpretive
  In some situation it is impossible to determine the nature of the relationship from a map. In such situation the cartographic text undertake the responsibility of presenting the relationships interpretation for users. The cartographic text reduces the necessity to go back to map legend or to do comprehensive measurement on the map (Major river, highway, state city).

3- Categorization

The categorization of theme is usually associated with a categorization of text symbols. The categorization of text serves the hieratical function of labelling.

4- Regulatory

Maps can be of legal, administrative, or political nature. In such kind of maps, more text is added to emphasize the territories, postcode zoon, and cadastral features.

5- Analytical

Cartographic text can go beyond depicting feature and can link the users with their interpretation based on the relationships among maps' features.

6- Confirmative

Text is used to confirm the meaning of the graphics that can be included in maps (Additionally to the legend text)

7- Determinative (Legend)

This function is used when the direct absolute measurement from the map is difficult, in such situation, table of actual data maybe used alongside. For example, a coropleth map which classifies data into broad categories or use non quantitative visual variable to depict quantitative data (the use of texture to depict values).

Text function associated with the complementary component of the map:

1- Temporal position

The cartographic text can refer to events happened in the past (e.g., historical battles) or events (or infrastructure) planned in the futures

2- Meta-data

This function refers to the nature of the source data. A completion diagram is commonly used as marginal information. Additionally, the relation to the data source can be referenced at the margin of the map.

3- Reference

Text on the mapface may explicitly refer to other materials which are:

- placed on the margins: as presenting the names of numbered symbols over the mapface to introduce their functions or to mention the labels when the mapface is too crowded with symbols;
- Presented in the form of supplementary but important information to use the map. In such case the cartographic text can be placed under map title or in the forefront of the legend;
- presented in additional publication to convey more about the map and its component.

Treating map texts as a symbol invites the cartographer to consider using the visual variables to present different characteristics of the cartographic text. The application of visual variables to cartographic texts is illustrated in the fourth column of Figure 1.1. The semiotic of cartographic text design follow the same rules of point, line and areal symbols (Tyner, 2010). Based on Withycombe's (1929) typographic essentials, the production of 'good' cartographic texts which provide a high level of communication, their visualization and thus their design should consider these principles:

- **perception and legibility**: the perception and legibility of cartographic text is controlled by its design. When compared to the legibility of normal (plain) text, cartographic text design was studied only in a limited manner. A few legibility rules were established to describe cartographic text functions (van den Worm, 2001; Slocum *et al.,* 2005; Kraak and Ormeling, 2010). However, fewer studies provided empirical evidence of the usability of cartographic text designs. In addition to the design itself, the display medium plays an important role in map readability and legibility with its resolution and size (van den Worm, 2001; Tyner, 2010);

- **balance and hierarchy**: hierarchy is one of the guided functions of cartographic text, which can present either equality of importance in the labelled features, when labels are harmonious in size, shape, etc.; or the hierarchy of importance in a set of features, when labels are degraded in case style, size, and combinations of variables.
- **discrimination**: the use of different visual variables (individually and combined) indicates more or less discriminable cartographic texts (MacEachren, 1995), which communicate differently with the map audience and the categorised map components. The various cartographic text shapes can be used to attract the eyes, while the use of unusual typefaces will make the reader stop and take notice (Tyner, 2010);
- **harmony**: similar to all elements on the map face, the cartographic text shall be organized in such a way that all categories of the text are compatible in their design and organization, i.e., using shapes that belong to similar families such as serif, historical typography, and case style;
- **classification**: the classification of cartographic text is correlated to feature classification because of the inherent link between them. This consideration is vital when labelling features that belong to the same class.
- **transferability**: when designing the cartographic text, it is important to consider how the typography will look on both the printed map and the digital map at different scales. Additionally, the link between the cartographic text sizes should be linked to the map scale in a way that preserves the proportion between type and feature dimensions.
- **availability, cost, convenience, and suitability for reproduction**: as with all map components, the availability of a typeface in the designing software shall be considered, along with how the typography can change when printing the map. Additionally, changes in the map's scale (with paper and digital maps) will affect the cartographic text legibility, therefore map display media shall be considered as a powerful controlling element of cartographic text design (Tyner, 2010).

The visualization of cartographic text should present the following requirements (Slocum *et al.,* 2005; Kraak and Ormeling, 2010; Tyner, 2010):

- cartographic they text should present nominal differences between different data classes;
- cartographic text should be able to address hierarchy. Thus, cartographic text visualization is associated with differentiating between more and less important objects or object categories; and
- cartographic text should relate to all feature types: point, line and area.

Cartographi text is added to complete the visualization functions and to deliver map messages. Its function is always associated with the attributed data. Typographic variables, such as shape, colour hue, texture directionality, and orientation, are effective means by which to attribute qualitative data. Size, colour value, colour saturation, texture size, and texture density are effective means by which to attribute quantitative data, as shown in Table 1.1.

1.1.4 Measuring a good label design

Robinson (1952) discussed in his book, *the look of a map*, that treating maps as art can lead to 'arbitrary capricious' decisions. To solve this issue he suggested either of two solutions. The first was standerizing maps (from data collection to all types of map use); as a consequence of standardizing maps, there would be no confusion or misunderstanding of the symbols' meanings. The second solution was to study and analyse the characteristics of map perception, so both symbolization and design decisions could be made based on the objective rules of map design. As he demonstrated, a map is both a piece of art and an organized system of data representation, but both treatments are required to enhance maps.

As cartographic text is one of the basic data types (Fairbairn, 1993; Tyner, 2010), the major concern when designing cartographic text is to provide cartographic products with high usability (Tyner, 2010). It was empirically proven by Hegarty *et al.* (2010) from their experimental study that a good design significantly facilitates user performance. The usability of a map is the capability of the map to be attractive, to be understood and to be used under specific conditions (Kraak and Ormeling, 2010). However, map usability is not a single, one-dimensional property of the map as an interface (Nielsen, 1993). Usability is usually determined by measuring effectiveness, efficiency and satisfaction (Faulkner, 2000; ISO, 1994). Before that, Nielsen (1993) measured usability in regard to five aspects:

- **learnability**; producing easy-to-learn maps is an internal goal that cartographers and cartographic researchers aim to achieve. For this, they consider design elements, map objects, and presentation methods (e.g. MacEachren, 1982). It must be considered that users normally use a map before they learn it. Therefore, complete map learnability is a time-consuming task. Cartographic text design shall be simple and obvious to enable quick learning. Complex design will disturb the learnability process;
- **efficiency**: fast map users can complete a task under certain conditions (e.g., Garlandini and Fabrikant, 2009). User efficiency could be related to map design, presentation conditions, digital versus paper maps, and other factors. Providing an efficient cartographic text design is a crucial aspect of enhancing map readability and interpretation (Slocum *et al., 2005)*;
- **memorability**: some map designs have proved to be more memorable than others. Reasons for this variation can be linked to the complexity of the design (e.g. MacEachren, 1982; Popelka and Brychtova, 2013) or the use of visual variable design elements themselves (Brewer, 1997). Some cartographic designs can have good memorability (like the artistic fonts), but this could distract from the design's functionality (Slocum *et al., 2005*);
- **errors**: a good map should have a low error rate for performing different tasks, and should influence users to correct their errors (Nivala, 2008). The distinct nature of cartographic text permits some types of error regarding its legibility;
- **satisfaction**; maps should be pleasant to use. For this purpose, map semiotics are devoted to presenting what is attractive and acceptable to users. For instance, coloured maps are preferred over monochromic maps. Some cartographic text designs can be preferred over others based on how visual variables are applied..

In both definitions of usability, the measurement falls into two categories: subjective and objective measurement. Therefore, to achieve good map design, both the subjective and the objective aspects of the design process are considered, including the cartographic text design. The boundaries between the subjective and objective design measurements are critically negotiable.

I. Objective measurements of labels' design

The objective measurement of design is determined by user performance and can be defined as the process of quantifying the efficiency of action (Gregory and Platts, 1995). Many methods can be used to express the user efficiency of a typographic design, including the following (Nielsen, 1993; Rubin and Chisnell, 2008; Barnum, 2010; and others):

- number or percentage of the correctly located targets;
- count of all incorrect target selections;
- time taken to complete a task; and
- eye tracing metrics.

Another method to determine user efficiency is to calculate the number of errors, incomplete tasks, and incorrect responses.

II. **Subjective measurements of labels' design**

The subjective design measurement is the evaluation process which reflects the designer's preferences for sets of multiple attributes (Thurston, 1991). With this definition, the evaluation is restricted to the designers, who can have different characteristics than users (Nielsen, 1993). Designers' and users' characteristics can deviate considerably based on the users' educational background, age, gender, experience, and many other characteristics. Therefore, preference for a label design is not the only measurement at the subjective level. The subjective level also includes measurements of characteristic variations between individual users. Therefore, the influence of user characteristics is considered to be subjective.

To address the characteristics of label design in terms of preferences, user preference can be measured in multiple ways (Rubin and Chisnell, 2008; Do Nascimento and Eades, 2008; Tyner, 2010; and others):

- the order of different designs;
- comparable choice;
- whether the design matches the expectations;
- whether the design functions appropriately for the required task;
- the ease of use, reading, learning, etc.;
- the familiarity of the design to the users (the aesthetic value of the design);
- the ease of correlation to the attributed object.

1.2 Rationale and synopsis

1.2.1 Research objectives and questions

The fact that we cannot read and interpret maps without their cartographic text (Fairbairn, 1993; Tyner, 2010), whose functions include delivering the map message, was the core motivation for the studies presented in this dissertation.

Although maps are symbolic in nature, maps cannot be fully communicated unless textual elements are provided. The textual elements are important for identifying map symbols by presenting a map legend, title(s), labels, and coordinates. Additionally, reading maps is not only a cognitive issue but also an issue related to map users' characteristics and their cultural background (Hambrick and Engle, 2002; Ooms *et al.,* 2012; Ooms *et al.,* 2013; and Popelka and Brychtova, 2013). This observation became a growing concern in the cartographic world in regard to providing the 'best' design, including the textual elements. Cartographic text design is a tripartite issue which requires consideration of the following aspects:

- the design of the textual element and how the design can differ in view of the cartographic text's function;
- user characteristics which can influence map perception, especially the influence of their level of experience and cartographic education, and the influence of gender differences on user performance; and
- lettering system differences and how these differences require distinctions in design rules for mono- and multi-lingual mapping.

In the context presented earlier, this dissertation comprises five research questions which are triggered by four research objectives:

- research objective A: improve the employment of visual variables in screen map design for better aesthetic and functional design;
- research objective B: investigate the influence of cartographic expertise and gender differences on map users' preferences and performance regarding cartographic text designs;
- research objective C: discuss the relationship between cartographic text colour and map background colour design; and
- research objective D: investigate the impact of lettering systems on cartographic text design.

The first research objective addresses the application of visual variables to the cartographic text. This objective is supported by the second objective, in which users determine both their preference for and the efficiency of the design. However, some previous cartographic work treated the first objective as the sole design issue, regardless of its link to the visual variables. Moreover, most (if not all) previous empirical studies on cartographic text were based on paper maps. The second research objective determines both user preferences and the performance of cartographic text designs in regard to user group characteristics, while the third research objective focuses on cultural differences among users. The fourth research objective defines the relationships between map backgrounds and labels in the foreground.

These four main research objectives are decoded into five, more specific research questions:

RQ 1: **How can the visual variables be applied to cartographic text, and more specifically map label?**

The automation age steered the attention of cartographers and map software programmers to think more about label placement than about how labels appear over the map face and what function they could serve. Therefore, in the last four decades, numerous studies of cartographic text placement have been presented, starting from Imhof's (1975) and continuing today. Cartographers have noted some labels' design criteria and design functions (Kraak and Ormeling, 2010; Slocum *et al.,* 2005; and MacEachren, 1995), considering them as a special sort of symbols. However, because they used typographic terminology, neither the link between this symbology and the visual variables nor the link between this symbology and typography for different languages was made.

***RQ2:* How do user characteristics influence the perception of label design?**

users' preferences express their acceptance of the typographic design and users' performance in response to specific stimuli defines their efficiency with the typographic designs (Nielsen, 1993; Rubin and Chisnell, 2008; Haklay and Nivala, 2010). Acceptance and efficiency as two parts of usability were studied with regard to differences in users' levels of expertise, cartographic education, and gender. Additionally, this aspect was considered when search strategies and how different users locate target labels were tested and analysed.

***RQ3:* How do lettering systems affect the usability of visual variables?**

Each lettering system has its own characteristics regarding word formation, letter assembly, typographic directionality, letter design (symbolic-pictorial) and many words' design issues (Huang, 2009; Chan and Lee, 2005; Hersch, 1994; Al-Harkan and Ramadan, 2005; Trenkle *et al.*, 2001; and others). These characteristics can be similar or dissimilar within a group of lettering systems. When applying the visual variables to different lettering systems, user preferences and performance of the produced designs can change due to the influence of visual variable characteristics on the typography of different languages.

***RQ4:* How do maps' background designs influence user preferences and performance together with the label design?**

Although cartographic design rules have been set for each map element individually, cartographers should not neglect the interactions between map elements, which can optimise users' perception of the map and improve their cognitive processes. Point, line, and areal features are usually visualized in a distinct manner, as each feature type has its own semiotics and syntax. However, little is known about the influence of map background colours on other elements in the foreground and especially on labels (MacEachren, 1995; Tyner, 2010).

***RQ5:* What type of search strategy do users follow when they search for labels over the map face and for certain map design conditions?**

An important impact for cartographic designers is the knowledge they can obtain about the way in which users look for targets on maps. This knowledge can improve the usability of the cartographic products in general, as it can identify some details of map perception. In addition to improving cartographic products, the search strategies can reveal many details related to the perception of cartographic designs.

Figure 1.3 illustrates how the research questions were distributed over the dissertations' chapters and the techniques and materials that were used to answer these questions. The inner circle represents the dissertation's chapters, divided into seven pie sections. The inner circle is followed by the methodology circle, which represents the method used for each chapter (questionnaire, time measurement, and eye tracking), which may overlap when more than one method was used in a certain chapter. The methodology circle is followed by the research question circle, which shows the research question(s) answered in each chapter. Finally, the outer circle represents the lettering system (Latin or multilingual) used in each corresponding chapter. Section 1.2.2 describes the content of the dissertations.

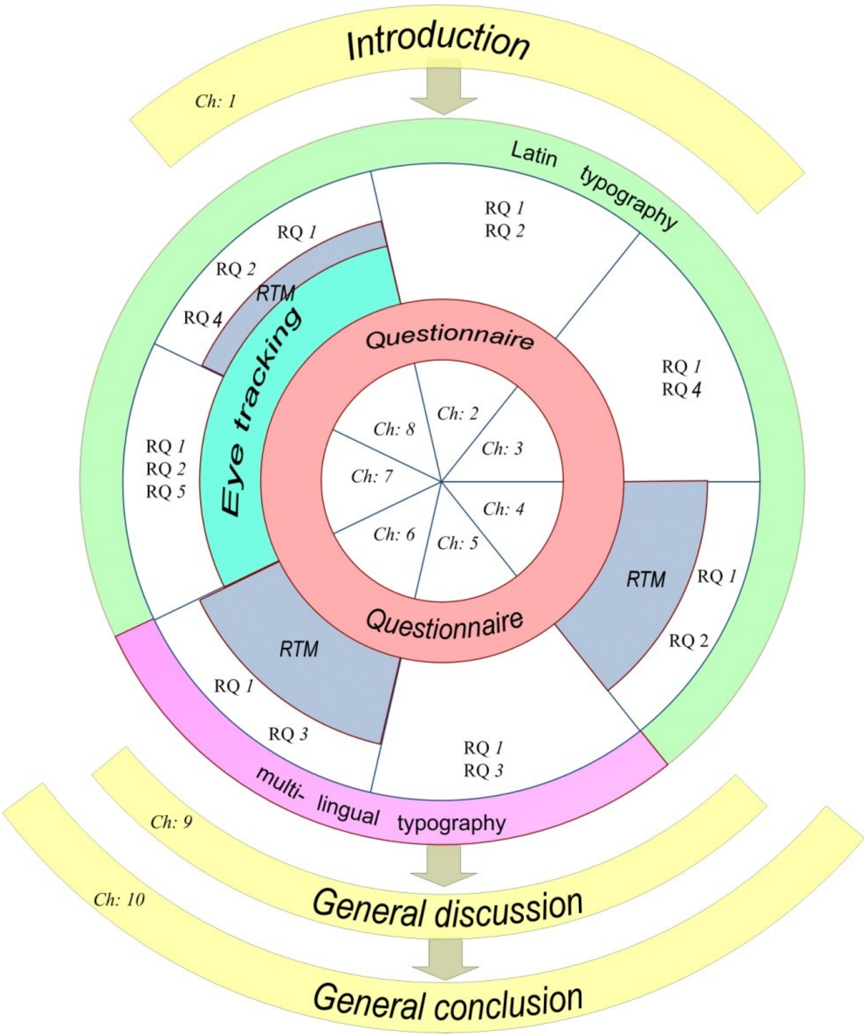


Figure 1.3 : Outline of the dissertation

1.2.2 Outlines of the dissertation

The dissertation describes a series of user studies which were constructed to answer the abovementioned research questions.

Chapters 2-8 correspond to articles that have been published or submitted to international peer-reviewed journals and conferences. A brief overview of these chapters – their relations, experimental design, and contents – is presented in Figure 1.4. The figure illustrates the technique used to answer each research question, the intersection and overlap of the chosen techniques, and the design of both the map and the typography. The basic root of these chapters is the application of Bertin's visual variables to cartographic text design.

Chapters 2-6 describe experiments whose typographic variables are size, shape, orientation, and texture. Meanwhile, Chapters 7 and 8 describe experiments dealing with the cartographic text colours. Screen maps were designed and implemented in these experiments, ranging from a basic design to complex topographic maps. Blank maps were used in the earliest stages of user testing; colours were gradually applied in backgrounds. Chapters 2, 3, 4, 7, and 8 determine user preferences and performance given Latin typographic designs, while in Chapters 5 and 6, Latin, Arabic, and Chinese typography are tested for cross-cultural comparisons of both efficiency and preference. Pre- or post-study questionnaires were used to obtain background information concerning participant characteristics. A detailed description of typographic and map designs, study techniques, and user characteristics can be found in the corresponding chapters as illustrated in Figure1.4.

Chapter 2: Cartographic Text Preference, Gender and Expertise Variation.

This chapter describes the application of Bertin's visual variables to cartographic text and the properties of various typographic designs. The preference of two user groups (gender and expertise) towards the cartographic design was determined and the variations between the sup-groups (females versus males, and experts versus novices) were analysed statistically. Few significant differences between the subgroups' preferences were identified. The study was published in the Cartographic Journal (Deeb *et al.*, 2012).

Chapter 3: Toward a Deeper Understanding of Cartographic Text Visualisation; Assessment of User Preferences and Colour Influence.

This chapter thoroughly investigates the influence of map background colours on user preference for the cartographic text design. Four sets of experiments were designed in parallel, each of which carries out the same typographic design, but with a basic change in the colour of the map backgrounds. The statistical analysis showed that no differences in user preferences resulted from changing the map background colour and thus there is no need to adapt the typographic design to the map colours. The study is accepted for publication in the Cartographic Journal (Deeb *et al.*, 2013b).

Chapter 4: Evaluating the Efficiency of Typographic Design; Gender and Expertise Variation.

Two user groups (gender and expertise) were involved in an experiment which used time measurement techniques to evaluate user efficiency. User efficiency for the cartographic text design was determined when users located target labels. Label designs have been adapted over the stimuli corresponding to the application of visual variables to the cartographic text. The study was published in the Cartographic Journal (Deeb *et al*., 2013a).

Chapter 5: Cartographic Text, the Preference of Label Design, and Cross-Cultural Comparisons.

After investigating user preferences for the cartographic text design of Latin lettering systems, cultural comparisons were made in two stages. In the first stage, user preference for Latin cartographic text was compared in parallel with Cyrillic cartographic text, as these lettering systems have many similarities in their word formatting characteristics. Second, user preferences for Latin lettering systems were compared with Arabic and Chinese lettering systems. The former comparison is between two similar lettering systems, while the latter comparison is among three distinguishable lettering systems. Both stages represented native users' preferences for the studied lettering system. The chapter describes the similarities and differences of map labelling in terms of user preferences for the application of visual variables to cartographic text and different lettering systems. The first part of the study was published in the proceedings of the 5th International Conference on cartography and GIS (Deeb *et al.*, 2014b). The second part is an accepted contribution to the 27th International Cartographic Conference (Deeb *et al.*, 2015c)

Chapter 6: Cartographic Text, the Efficiency of Label Design, and Cross-Cultural Comparisons.

In a the way in which users search maps is crucial to understanding their cognitive processes and is very important for conveying perception issues to map designers. Therefore, an eye tracking experiment was conducted to demonstrate users' search strategies and specifically to look for clusters of strategies. This study integrated several open source software packages to analyse scan paths. Chapter 7 continues to analyse the differences between user groups, as it distinguishes between the performance of experts and novices. The study also distinguished between females' and males' responses. Few clusters appeared in the demonstrated search patterns. This study is submitted to the ICC 2015 (Deeb *et al.*, 2015a).

Chapter 7: Assessment of Map Users' Search Strategies Using Multi-Package Software for Scanpath Analysis.

Understanding the way in which users' search on maps is crucial to understand their cognitive process, and is very important for map designer to convey perception issues. Therefore, an eye tracking experiment was established to demonstrate users' search strategies and specially to look for strategies' clusters. This study engaged several open source software to analyse scanpaths. Chapter 7 continues to analyse the differences between users' groups as it distinguished between the performance of expert and novices in the one hand. On the other hand the study also distinguished between females and males responses. Search patters were demonstrated where few clusters appeared. This study is an accepted contribution to the 27th International Cartographic Conference (Deeb *et al.*, 2015b).

Chapter 8: Background and Foreground Interaction, and the Influence of Complementary Colours on Search Tasks.

Visual variable colour is the focal point of this study. Coloured labels in the foreground were studied in comparison to black labels. Additionally, the interaction between colours in the foreground and their complementary colours in the background was studied. Eye tracking data were used to interpret user performance. As in Chapter 2, Chapter 4, and Chapter 7, Chapter 8 investigates the performance differences between both expertise and gender groups. This study was published in the Color Research and Application Journal (Deeb *et al*., 2014b).

Chapter 9: General discussion.

In this chapter, answers to the research questions are presented. A general discussion on the outcomes of the dissertation is included. Additionally, application boundaries, challenges, limitations, and future implications are examined.

Chapter 10: General conclusion.

This chapter presents a general conclusion on the dissertation results.

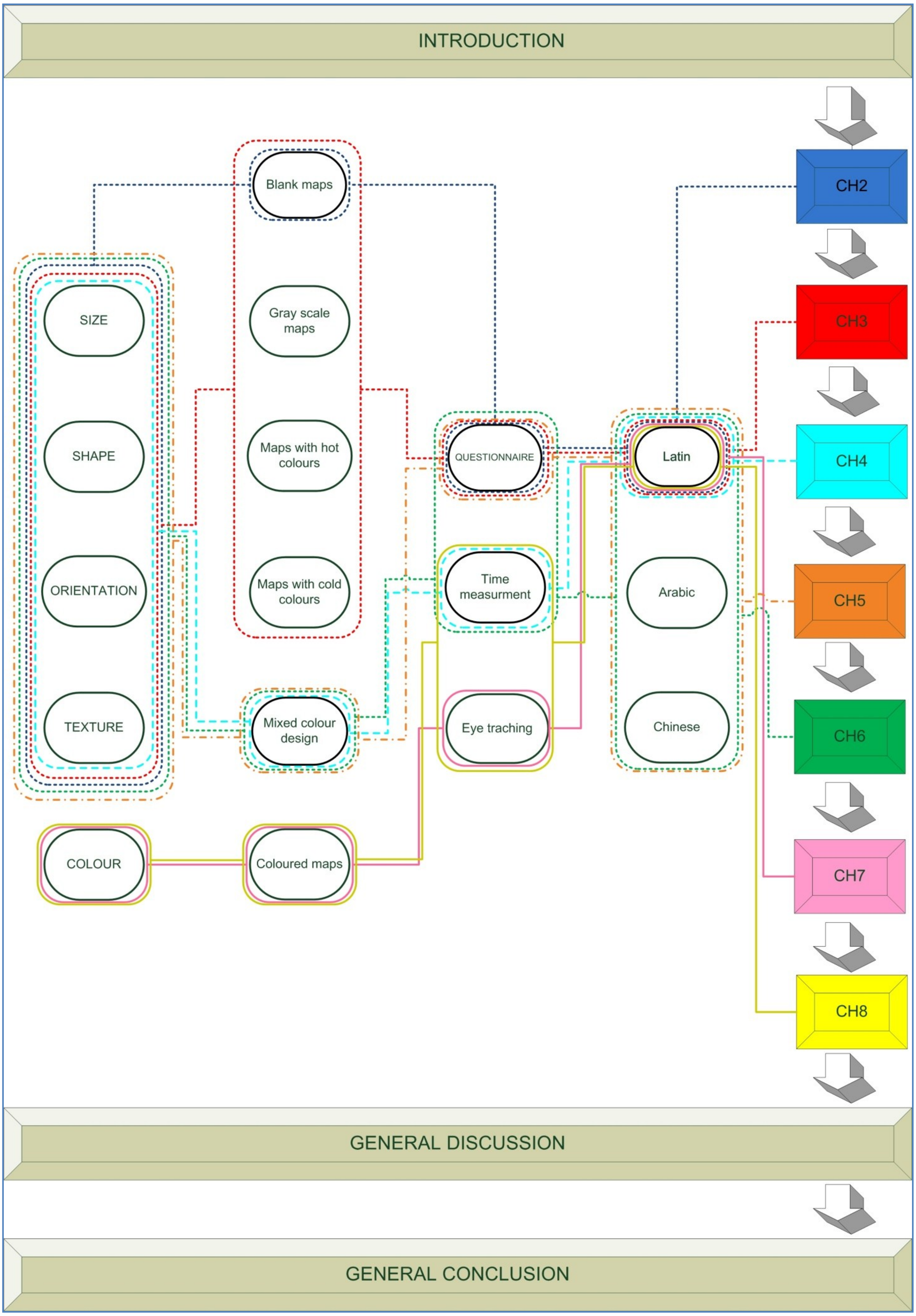


Figure 1.4: The relationships between chapters of the dissertation

Figure 1.3 and Figure 1.4 illustrate how the study of textual visual variables is presented over the dissertation's chapters and the method used for each chapter, in addition to the different lettering systems. Figure 1.4 shows that there is no constant relationship between chapter components. A one-to-one relationship between the visual variables and other dissertation elements is found in Chapter 7; a one-to-many relationship is found in Chapter 8; a many-to-one relationship is found in Chapter 2; and a many-to-many relationship is found in Chapters 3, 4, 5, and, 6.

1.2.3 Out of scope

The focal point of this research is the improvement of cartographic text design on screen maps. Because cartographic text design was not discussed thoroughly in previous research, many research questions arose while defining the objectives of this research. We had to set some boundaries for each research question. Therefore, the definition of each research question includes limitations which are discussed within each chapter.

Because only a few studies have considered cartographic text design and treated text semiology as equal to point, line, and area semiology, we had to make some critical decisions for each research step to provide a framework for the dissertation. As this dissertation cannot fully cover all aspects of cartographic text design, limitations were presented in five major points:

I. **labels as a sample of cartographic text:**

Fairbairn (1993) defined fifteen classes of cartographic text, which are associated with the major functions of cartographic text. The dissertation covers the functions that can be provided by labels. Although we refer to labels as 'cartographic text,' we must note that the text of legends, coordinates, marginal information, and metadata were not included in this study;

II. **labelling linear data:**

The research encountered label design issues in both the labelling of point data on topographic maps and the labelling of areal data on thematic maps. Labelling linear data on either topographic or thematic maps is not discussed in this dissertation, as it introduces label design issues regardless label placement implications;

III. **desktop screen presentation:**

In 1967, Bertin specified paper maps as the only map presentation vehicle to which his variables are applicable. Our research investigated the applicability of these variables to desktop screen maps. However, many display vehicles were not discussed in the research, such as small smart phones and GPS routing devices (with screens of only a few inches), tablets and notebooks (approximately 10 inches), and larger-sized screen , such as monitors, flat screens, touch tables and television screens (20-50 inches). Additionally, paper maps of all printing sizes were not tested in the included experimental studies;

IV. **3D cartographic data labelling:**

3D visualization adds an extra dimension to the maps and adds similarly third dimension to the included text. Therefore, we cannot predict how map users will respond to cartographic text designs represented in 3D based on their performance on 2D maps;

V. **interactive maps:**

Both user preference and user efficiency were tested with static screen maps. The influence of interactive operations (such as zoom, panning, and changing layers) was not specifically tested;

VI. **Effectiveness:**

The dissertation focused on two aspects of map usability. Both user preference for the cartographic text designs and user efficiency with these designs were studied. The dissertation did not fully cover the third aspects of usability (effectiveness). Only the effectiveness of colour design was considered.

The effectiveness of cartographic texts design refers to the design ability to do what it should do (Rubin and Chisnell, 2008), and since the link between the functions of cartographic texts and their design was not the main objective of this dissertation, effectiveness of cartographic text design was not discussed thoroughly over the dissertation's chapters.

In addition to what is mentioned above, it must be considered that only two user characteristics were studied and analysed: expertise and gender.

1.3 Reference

Al-Harkan, I. M. & Ramadan, M. Z. 2005. Effects of Pixel shape and color, and matrix pixel density of Arabic Digital Typeface on characters' legibility, International Journal of Industrial Ergonomics, 35(7), pp.625-664.

Arntson, A. E. 2011. Graphic design basics. 5th edition, Thomson Wadsworth, Boston.

Barnum, C. M. 2010. Usability testing essentials: ready, set... test! Morgan Kaufmann, Burlington.

Bartz Petchenik, B. 1974. A verbal approach to characterizing the look of maps. The American Cartographer, 1(1), pp.63-71.

Bertin, J.1967. Semiology of Graphics: Diagram, Networks, Maps, Translated by William J. Berg, 2011, Esri press, California.

Board, C. 1972. Cartographic communication. Cartographica: The International Journal for Geographic Information and Geovisualization, 18(2), pp.42-78.

Board, C. & Taylor, R. M. 1977. Perception and maps: Human factors in map design and interpretation. Transactions of the Institute of British Geographers, pp.19-36.

Brewer, C. A., MacEachren, A. M., Pickle, L. W., & Herrmann, D. 1997. Mapping Mortality: Evaluating Colour Schemes for Choropleth Maps, Annals of the Association of American Geographer, 87(3): pp.411-438.

Caivano, J. L. 1990. Visual texture as a semiotic system, Semiotica, 80(3/4), pp.239-252.

Caivano, J. L. 1991. Cesia: A System of Visual Signs Complementing Color, COLOUR research and application, 23(3), pp.258-268.

Caivano, J. L. 1998. Color & Semiotic: a two way street, COLOUR research and application, 23(3), pp. 390-401.

Cambridge online dictionary, http://dictionary.cambridge.org/, used on 14/10/2014

Chan, A. H. S. & Lee, P. S. K. 2005. Effect of display factors on Chinese reading times, comprehension scores and preferences, Behaviour & Information Technology; 24(2), pp.81-91.

Deeb, R., Kurban, A., & De Maeyer, P. 2015c. From Monolingual to Multilingual Mapping: What is users' preference?. 27th International Cartographic Conference, Rio de Janeiro. Accepted – ICC 2015.

Deeb, R., Ooms, K., Popelka, S., Van Etvelde, V. & De Maeyer, P. 2015b. Assessment of map users' search strategies using multi-packages software for scanpaths analysis. 27th International Cartographic Conference, Rio de Janeiro. Accepted – ICC 2015.

Deeb, R., Ooms, K, Kurban, A., Ablikim, A., & De Maeyer, P. 2015a. Cartographic text, the efficiency of label design and cross culture comparisons. Information Visualization. Submitted 2015.

Deeb, R., Ooms, K., Van Etvelde, V. & De Maeyer, P., 2014b. Background and foreground interaction: complementary colours' influence on search task. Color Research and Applications. Accepted 28 August 2014

Deeb, R., Bonchev, S., Bandrova. T.,& De Maeyer, P., 2014a. How to design users' friendly multilingual maps? Cyrillic and Latin Labels. 5th International Conference on cartography and GIS. Riviera

Deeb, R., Ooms, K., Van Etvelde, V. & De Maeyer, P., 2013b. Toward a deeper understanding of cartographic text visualisation: assessment of user preferences and colour influence. The Cartographic Journal. DOI: http://dx.doi.org/10.1179/1743277413Y.0000000073

Deeb, R., Ooms, K., Vanopbroeke, V. & De Maeyer, P., 2013a. Evaluating the Efficiency of Typographic Design: Gender and Expertise Variation. The Cartographic Journal. 51(1), pp. 75-86.

Deeb, R., Ooms, K. & De Maeyer, P.,2012. Typography in the Eyes of Bertin, Gender and Expertise Variation. The Cartographic Journal, 49(2), pp. 167-185.

Denegre, J. 2005. Sémiologie et conception cartographique, Hermes-science, Paris.

DiBiase, D., MacEachren, A. M, Krygier, j. b. & Reeves, C. 1992. Animation and the role of map design in scientific visualisation. Cartography and geographic information science, 19 (4), pp.201-2014.

Do Nascimento, H. A., & Eades, P. 2008. User Hints for map labelling. Journal of Visual Languages & Computing, 19(1), pp.39-74.

Dransch, D. 1995. Temporale und nontemporale Coumputer-Animation in der Kartographie , 15. Selbstverlag Fachbereich Geowissenschaften, Freie Universität Berlin.

Fairbairn, D.J. 1993. On The Nature of Cartographic Text, The Cartographic Journal, 30(2), pp.104-111.

Faulkner, X. 2000 Usability Engineering, NY, Palgrave.

Garlandini, S., & Fabrikant, S.I. 2009. Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization'. In: Lecture Notes in Computer Science, Spatial Information Theory, Hornsby, K.S. *et al.* (eds), 5756, pp.195-211, Springer-Verslag, Berlin Heidelberg.

Gregory, A. N. M. & Platts, K. 1995. Performance measurement system design: A literature review and research agenda. International Journal of Operations and Production Management, 15(4), pp.80-116.

Haklay, M. M. & Nivala, A. M. 2010. 'User-centred Design', In: Interacting with Geospatial Technologies. ed. by Haklay, M. M. chapter 5,pp. 91-106, John Wiley& Sons. London.

Hambrick, D. Z. & Engle, R. W. 2002. Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. Cognitive psychology, 44(4), pp.339-387.

Hegarty, M., Canham, M., & Fabrikant, S. I. 2010. Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. Journal of Experimental Psychology: Learning, Memory, and Cognition, 36(1), pp. 37-53.

Hersch, R. D.1994, Font rasterization: the state of the art. In From object modelling to advanced visual communication, pp.274-296. Berlin Heidelberg.

Hinks, M. 2009. Perceptual and design principle for effective and interactive visualizations, In: Trend in Interactive Visualization, ed. by State-of-the Art Survey. Eds: Seinstra, E. Z., Andriaansen, T., and van liere, R., Chapter 7, pp.155-174, Springer, London.

Huang, D-L., Rau, P-L. P. & Liu, Y. 2009. Effects of font size, display resolution and task type on reading Chinese fonts from mobile devices. International Journal of Industrial Ergonomics, 39(1), pp.81-89.

Imhof, E. 1975. Positioning Names on Maps, The American Cartographer, 2(2). pp.128-144.

ISO (1994). ISO9241-11: Ergonomic requirements for office work with visual display terminals (VDTs), Part11: Guidance on Usability.

Köbben, B. & Yamman, M. 1996. Evaluating dynamic visual variables. In Proceedings of the seminar on teaching animated cartography, ACI/ICA, Madrid, pp.45-51.

Koch, W. G. 2001. Jacques Bertin's theory of graphics and its development and influence on multimedia cartography, Information Design Journal, 10 (1), pp.37–43.

Kolácný, A. 1969. Cartographic information - A fundamental concept and term in modern cartography. The Cartographic Journal, 6(1), pp.47-49.

Kornfeld, A. L., Schiewe, J. & Dykes, J. 2011. Audio cartography: visual encoding of acoustic parameters. In Advances in Cartography and GIScience. Volume 1. pp.13-31. Springer Berlin Heidelberg

Kraak, M- J. 2001. Settings and Needs for Web Cartography, In: Web Cartography, Kraak, M.J. and Brown, A. (eds), Chapter 1, pp. 1-7, Taylor & Francis, London.

Kraak, M. J. & Ormeling, F. 2010. Cartography. Visualization of Spatial Data, 3rd edition, Person Prentic Hall, Harlow.

MacEachren, A. M. 1982 .The Role of Complexity and Sympolization Method in Thematic Map Effectiveness., Annals of the association of American Geographers, 72(4), pp.459-513.

MacEachren, A. M. & Kraak, M. J. 1997. Exploratory Cartographic Visualization: Advancing the Agenda. Computer and Geoscience, 23 (4), pp.335-343.

MacEachren, A. M & Taylor. D. R. F. 1994. Visualization in Modern Cartography, Pergamon, Oxford.

MacEachren. A.M. 1995. How Maps Work, Representation, Visualization and Design, the Guilford press, NY.

Moore, P. & Fitz, C. 1993. Using Gestalt theory to teach document design and graphics, Technical communication quarterly, 2(4), pp.389-410.

Morrison, J. L. 1974. A Theoretical Framework for Cartographic Generlization with Emphasis on the Process of Symbolization. International Yearbook of Cartography, 14, pp.115-127.

Morrison, J. L. 1984. Applied Cartographic Communication: Map Symbolization For Atlases,Cartographica, 21(1), pp.44-84.

Nielsen, J. 1993. Usability Engineering, Academia Press, San Francisco.

Nivala, A. M., Brewster, s. & Sarjakoski, T. 2008. Usability Evaluation of Web Mapping Sites. The Cartographic Journal, 45 (2), pp.129-138.

Ooms, K., De Maeyer, P., Fack, V., Van Assche, E., & Witlox, F. 2012. Interpreting maps through the eyes of expert and novice users. International Journal of Geographical Information Science, 26(10), pp.1773-1788.

Ooms, K., De Maeyer, P. & Fack, V. In Press. Listen to the Map User: Cognition, Memory, and Expertise.The Cartographic Journal. DOI: http://dx.doi.org/10.1179/1743277413Y.0000000068

Oxford online dictionary, http://www.oxforddictionaries.com/, on 14/10/2014

Poidevin, D. 1999. La carte, moyen d'action. Ellipses, Paris.

Popelka, S., & Brychtova, A. 2013. Eye-tracking study on different perception of 2D and 3D terrain visualisation. The Cartographic Journal, 50(3), pp. 240-246.

Robinson, A. H. 1952. The look of maps; An examination of cartographic design. Madison: University of Wisconsin Press.

Robinson, A. H., & Petchenik, B. B. 1975. The map as a communication system. The Cartographic Journal, 12(1), pp.7-15.

Rubin, J. & Chisnell, D. 2008. Handbook of Usability Testing. How to Plan, Design and Conduct Effective Tests, 2nd edition, Wiley Publishing, Indianapolis.

Shannon, C. E. 2001. A mathematical theory of communication. ACM SIGMOBILE Mobile Computing and Communications Review, 5(1), pp.3-55.

Slocum, T. R., McMaster, R. B, Kessler, E. C. & Haward. H. H. 2005. Thematic Cartography and Geographic Visualization, 2nd edition, Person Prentic Hall, Upper Saddle River, NJ.

Thurston, D. L. 1991. A formal method for subjective design evaluation with multiple attributes. Research in Engineering Design, 3 (2), pp.105-122

Trenkle, J., Gillies, A., Erlandson, E., Schlosser, S., & Cavin, S. 2001. Advanced in Arabic Text Recognition. In: the proceeding of symposium on Document Image Understanding Technology , pp.158-168, Columbia: Maryland

Tyner, J. A. 2010. Principles of Map Design, The Guilford press, NY.

Uttal, W. R. 1988. On Seeing Forms. Hillsdal, NJ:Erlbaum.

van den Worm, J. 2001. 'Web Map Design in Practice', In: Web Cartography, ed. by Kraak, M.J. and Brown, A. Chapter 7, pp. 87-107, Taylor & Francis, London.

Ware, C. 2004. Information Visualization, Perception for Design, 2nd edition, Elsevier, China.

Withycombe, J. G. 1929. Lettering on maps. The Geographical Journal, 73(5), pp.429-435.

2

Cartographic Text Preference, Gender and Expertise Variation

Modified from :Deeb, R., Ooms, K. & De Maeyer, P.,2012. Typography in the Eyes of Bertin, Gender and Expertise Variation. The Cartographic Journal, 49(2), pp. 167-185.

Preface

The main goal of this Chapter is to set a group of typographic criteria to suit a wide array of map users. A group of Bertin’s visual variables were applied individually and associatively for the same set of labels. Two kinds of maps (with point and areal objects) were presented to expert and novice map users and analysed accordingly. Additionally, the effect of gender variation was taken into account. The data were aggregated and studied for each graphical variable. For some combinations of Bertin’s variables, statistically significant differences were detected in the preferences of the different map users (e.g., female versus male and novice versus expert). Consequently, we identified which graphical variables (individually or combined) were more preferred by specific user groups in relation to their application on cartographic text.

2.1 Background and Objectives

Reading a map is a task that cannot be accomplished unless textual elements are provided. These texts, and more specifically the labels on the map image, can be considered the fourth symbol type in addition to points, lines and areas. This finding implies that certain rules should be applied in the typographic design (Fairbairn, 1993).

Board and Taylor (1976) demonstrated the advantages of using a conventional symbology and its influence on reading names. Gerber (1981) identified reading labels as a midterm level of the reading process, while Robinson and Petchenik (1976) defined the reading as an active process and a transaction between the individual and the environment. Therefore, it is important to know how much information is transferred to each reader.

Variations in how labels are presented on the map (e.g., size and colour value to indicate a hierarchy) facilitate the interpretation of the contents (Imhof, 1975). Bartz (1970a) also examined the influence of font size and its adequate dimension. When used incorrectly, however, the variations in label presentation may influence the interpretation of the map's contents in a negative way. Therefore, it is necessary to know which type of variation can be used on map labels to allow efficient interpretation.

Bertin (1967) defined six visual variables (size, value, texture, colour, orientation and shape) that are applied to the symbols on a map to visualise properties of objects or links between objects, including associativity, selectivity, order and quantity. In Table 2.1 the links between the six visual variables and their properties are illustrated.

Table 2.1: Levels of organization of the visual variables. After Bertin (1967, fig. 2, p.69)

Visual variable	***Associativity***	***Selectivity***	***Order***	***Quantity***
Size		■	■	■
Value		■	■	
Texture	■	■	■	
Colour	■	■		
Orientation	■	■		
Shape	■			

Text can be considered the fourth symbol type, therefore these visual variables should also be applicable in this case (Fairbairn, 1993). Text objects on a map do not have the same design rules as point, line and area objects. It is distinct in the way visual variables are applied. The 'translation' of the application of these visual variables to text symbol type is described below and illustrated in Table 2.2:

- **size** is applied to text by changing the font size, consuming more space on the map for larger corpus and less space for smaller corpus. In this article, points (dots) are used to indicate texts sizes. Boldness was added to this variable as the word consumes more space for the same font size. Size provides selective, ordered and quantitative perception. Therefore, variations in size can indicate quantitative information like the population of a city, but this can also be used to order information (discrete classes indicating population levels) using levels of hierarchy among the different label classes;
- **value** is defined by Bertin (1967) as the ratio between black and white perceived on a given surface. Changing the value of the typographical variable serves order and selectivity, but it

is still a critical concern when the printing method is used (grey tone or coloured), in addition to the background of the map. Based on the variation in value different label levels with hierarchal purposes are served;

- **texture** is represented by a combination of differing text characteristics. The texture of typographical symbols is linked with the overall layout of the individual word, such as the narrow or wide spacing of letters. In addition to that, texture is recognized by the overall view of all textual elements on the map image, reflecting the relationship of labelled features. Size and value work together to provide infinitive array of textures. Texture is an associative variable that allows users to identify variations and group correspondences within all depicted categories. Selectivity and order can also be visualised using this variable. These properties can thus be used to indicate a distinct themes (e.g. continent versus ocean) or certain label hierarchy (e.g. city versus village);
- **colour** (or hue) of texts is the attribute of the visual sensation on a fixed saturation and value. Colours provide both an associative and a selective perception. Text colour can work functionally and is usually employed to differentiate between themes, such as blue for water bodies and red for danger;
- **orientation** can be applied to text either on each individual character (italic), or by tilting the whole word. Orientation plays an important role in the associative and selective perception. Unlike in books, texts on maps can follow many orientations to stress the shape of the labelled object, such as rivers or countries;
- **shape** is represented using different fonts. 'The world of shape is infinite' Bertin said, and this holds also true for text objects with an unlimited number of fonts. Associativity is the property that is obtained by using different shapes, which allows the map reader to distinguish between different categories and group similar objects. Consequently, shape can describe different themes at the same level of importance. This is commonly used when labelling, for example, land use maps.

Table 2.2: Bertin's variables on text

Variable	Examples		
Size	Cartography	Cartography	Cartography
	Cartography	**Cartography**	**Cartography**
Value	Cartography	Cartography	Cartography
Texture within a word	Cartography	Cartography	Cartography
Texture in a group of words	CARTOGRAPHY Cartography cartography	CARTOGRAPHY Cartography *cartography*	CARTOGRAPHY *Cartography* Cartography
Colour	Cartography(red)	Cartography(green)	Cartography(blue)
Orientation	Cartography	*Cartography*	Cartography
Shape	Cartography	Cartography	Cartography

After Bertin defined these visual variables and their properties, some authors took a closer look at them in the light of new psychological theories or cartographic development. Imhof (1975) added the variable position which he also applied to texts as a label can be placed at a number of (fixed) locations relative to its associated objects. He described the four preference positions of labels associated with point objects as the upper right corner, lower right corner, upper left corner and lower left corner. The investigation of the optimal position for labels, however, this concern is not discussed in this chapter. MacEachren (1995) thoroughly discussed the visual variables based on their X, Y dimensions together with their numerical, ordinal and nominal applications. In addition to the dimensions of the symbology, modern cartography also allowed the development of new variables, such as sound (MacEachren, 1994) and time (Kraak *et al.*, 1997). Koch (2001) referred to Bertin's variables as an application of the Gestalt laws, which describe rules of similarity and good design. These laws have been extended after Bertin, but they are still valid in the modern cartography. For $2^{1}/_{2}$ D and 3D representations Slocum *et al.* (2005) rephrased Bertin's 2D variables using spacing, size, perspective height, orientation, shape, arrangement and finally hue, lightness and saturation.

The legibility of maps has been a concern of cartographers since they started developing cartographic products. Additionally, maps are widely used products by many different groups of users. Consequently, text as a distinct component requires special care in the long-term development of maps. Locating names based on typographic variations and measuring the completion times of tasks was explored by Bartz (1970b), who conducted a series of experiments in which participants had to locate names on a map. This gave insights in the efficiency of certain typographic variables towards the map user. Foster and Kirkland (1971) studied the association between text and different colours. Phillips *et al.* (1977) compared the response times of users related to different text characteristics, such as size, boldness and shape, for difficult and easy names. Kraak and Ormeling (2010) defined rules for texts on maps to improve their readability. They also suggested utilising different textual variables to produce a more readable map that also meets the needs of text functions.

The studies mentioned above investigated the legibility of labels using the amount of time a map user needs to complete a task. However, none of these studies considered the combination of all Bertin's variables and their implementation on the textual information on maps as well as their influence on the legibility of the map. These studies also used paper maps to examine a range of individual text variables and fixation time measurement to indicate the 'best' text visualisation. Furthermore, no link was made from the application of Bertin's variables to texts. This chapter aims to extend these studies by measuring user preferences when applying Bertin's variables to texts (excluding colour and value), taking into account experience and gender as each of these groups may respond differently to maps.

When designing a map, different choices have to be made: type of map, theme of the map, map audience, use of the map, etc. Two of the most important types of maps are topographic and thematic maps. The way the labels are placed and visualised on these maps greatly depends on their type, theme and audience. In the study described in this chapter, two basic map types are used: maps with point data and maps with areal data which are all labelled. The first category (with point data) can refer to topographic maps, isopleths maps and dot maps whereas the second category (areal data) can refer to the majority of qualitative and quantitative thematical maps such as choropleth Maps, soil maps, socio-cultural maps and socio-economical maps. A blank background maps are used as a basic of extensive studies where more complex map design will be tested and colour and value will be tackled.

The experiment is a user study in which the participant must choose one of two maps that have variations in how labels are visualised (cf. Bertin's variables). Since different types of users are included in the study (females, males, novices, expert), the outcome of the study reveals whether cartographic training and practice influences label preference on maps. This method allows differentiating the variables used on maps according to who the map intended for. The objective of the experiment is to answer the following questions:

- What size does the audience prefer in its ratio to the general space of the map?
- What is the influence of bold fonts on the audience's preference?
- What kind of fonts does the user tend to accept more?
- Which aspect of the orientation variable do the audience prefer the most?
- What is the influence of simple label textures and complex label textures on users' preference?
- Do all users have the same preferences when there are variations in the level of experience in map use and in gender?
- How different levels of label hierarchy can be presented using Bertin's variables and what is the impact of these designs on users' preference?

2.2 The Methodology

An experiment was constructed to examine the users' preference towards the application of Bertin's visual variables to texts for both point and areal data. The study provides an overview of map users' preferences of labelling characteristics based on variation in experience and gender.

2.2.1 Participants

A group of 80 map users participated in the study. They were divided according to the criteria experience and gender. The novice group included 50 participants in the beginning of their geographical education, with no previous training in cartography. The other group of participants consisted of 30 experts who work with maps on a daily basis and have at least a master's degree in geography. Of the 80 participants, 35 were female and 45 male divided in balance across experts (15 females and 15 males) whereas the novice group consisted of 20 females and 30 males. The average age of the users was 23.3 years, with the experts' average on 27.3 years and the novices on 20.4 years. By taking into account experience and gender, it was possible to detect what influence the users' backgrounds had on their preferences regarding labels on maps.

2.2.2 Task, stimuli and data

Forty-one maps were presented to the users during the experiment in pairs or triples, forming 79 sets of questions. Base maps were drawn at the scale of 1:10 000. These maps were designed with a blank background in order to acquire neutral results of users' preference. The study had both a between- and within-user design. The first part of the experiment involved maps populated with point data and their associated labels. The graphical variables size, shape and texture were applied to the map labels to visualise levels of importance in the labels and thus in the associated objects. An example of such a pair of maps is depicted in Figure 2.1. During the second part of the experiment, variations in the size, shape and orientation of labels associated with areal objects were investigated. Two maps from these experiments are illustrated in Figure 2.2. Bertin's variables were applied (individually and combined) and integrated in a coherent structure for both series. Colour and value were excluded from these stimuli because the function of these variables is mainly to indicate the nature of the visualized object (Kraak and Ormeling, 2010) which is not the domain of this research. In addition to that, text colour and value cannot work functionally with a blank background because of the consistent interaction between the text foreground and the background's colours.

To avoid biases in the answers due to resolution and size differences, all participants completed their test on a flat screen with a 1280x1024 resolution. Each participant followed the same order of maps in a sequence that lasted 20-30 minutes.

These trials were embedded in an online questionnaire in which two or three maps appeared on the screen simultaneously. The participant then indicated a ranking, which was subsequently stored in a database. The result was an ordered list of the maps for each participant, indicating his preferences on the readability of the maps for the different visualised graphical variables.

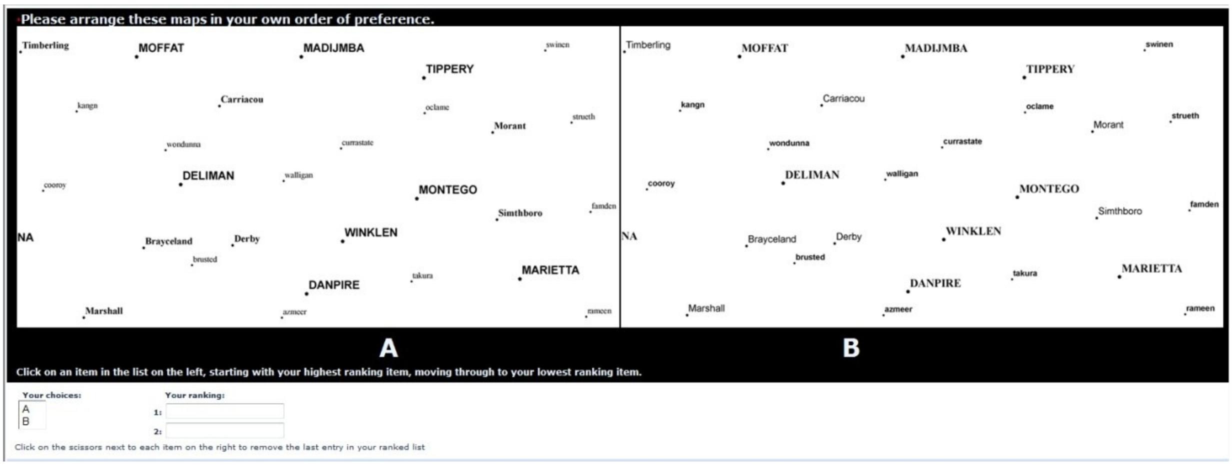


Figure 2.1: Example of the ranking task for point data (variable: texture)

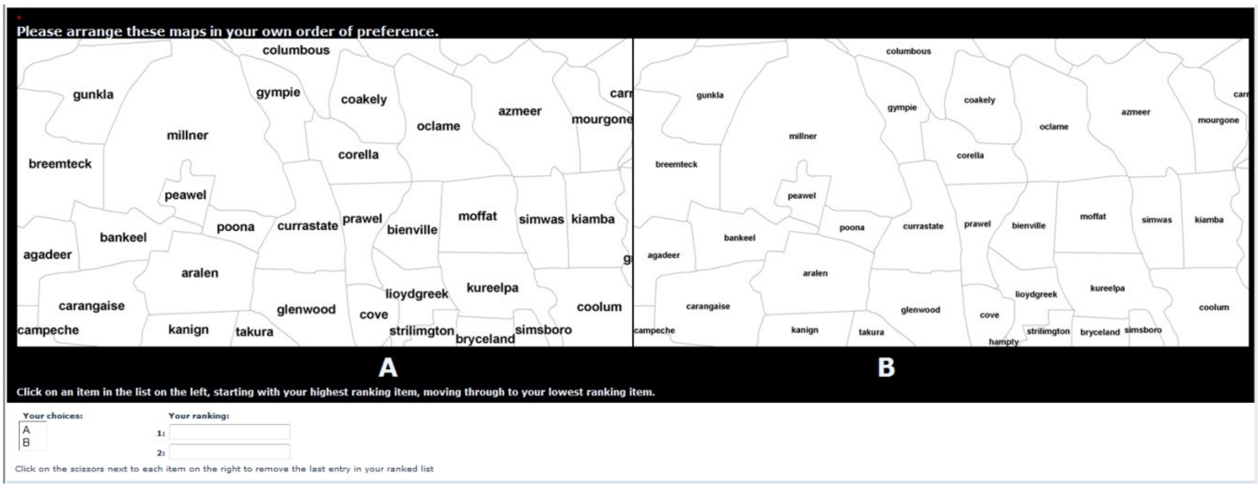


Figure 2.2: Example of the ranking task for areal data (variable: size)

2.3 Results and Discussion

2.3.1 Measurement of preference and typographic quality for all users

Map aesthetics is an important aspect of cartographic work. To recall organised cartographic symbols factually and sensibly, care should be taken in choosing the adequate symbols for a wide range of audiences. Therefore, a variety of typographic symbols were presented to the map audience to test which presentation was preferred. The collected data were accumulated and restructured to indicate the preferable criteria between each set of maps.

A selection of criteria was used to obtain the most preferred quality that measures aesthesis. Typographic criteria were analytically organised and grouped in order to reflect four major mapping purposes:

- the equivalent use of criteria of names and labels;
- the hierarchal use of criteria of names and labels;
- the thematical use of criteria;
- a combination of the above purposes, for example to emphasize certain elements.

Thanks to the combination possibilities of visual variables, numerous functionality can be expressed using different level of complexity. One of the most important functions is the hierarchical importance which can be embedded in three level of complexity. Complexity refers to the systematic combination of the visual variables. Accordingly, the labels (and maps) used in this study were divided into three levels of complexity:

- first level of complexity: one of Bertin's visual variable was assigned to labels and text (e.g showing hierarchy using only different font sizes);
- second level of complexity: two variables were combined or labels were presented in a hierarchical way using a visual variable in addition to font size;
- third level of complexity: formed by using more than two associated variables. Complex hierarchy is formed by adding tow visual variables to font size which can be link to it to form different hierarchal textures (case styles, boldness, etc.).

2.3.1.1 Size

Regarding the first level of complexity, size as a typographical variable varied from 8- to 14-font size. The sizes of the depicted labels were controlled according to the test structure and map's dimensions. These sizes were compared pair wise and summed to reveal which font size was preferred. The results, listed in Table 2.3, show the order of preference to be 10-8-12-14. Although test structure defined the used map scale, the users' preference was calculated for the overall view of map and the ratio of label size to the area size. However, when only considering the labels' font sizes, the largest size (14 point) showed the least interest from the users and consequently, it did not acquire a high rate of aesthetics evaluation. Here, the label size (14 point) is considered too big relative to area they cover. Size 12 point is also rather large but it has a noticeable higher preference rate than 14 point. The smallest size did not match the highest preference rate compared to size 10 point which is the most preferred size. Consequently, this reflects the highest aesthetics design by the symmetry and harmony between the label font size and the district size. This result demonstrates the relation between the labels to their depicted area.

Table 2.3: Size overall preference

Size	8	10	12	14	total
8	0	33	49	62	144
10	47	0	60	74	181
12	20	31	0	71	122
14	9	6	18	0	33

The second level of complexity was also related to the size of the labels; however, in this case, bold and uppercase characters were used to indicate hierarchy. The variations in the case styles were presented using ‘all letters in lowercase’, ‘only first letter in uppercase’, ‘all letters in upper case’. These levels of hierarchy were both tested with a serif (Times New Roman, TNR) and sans serif (Arial) font. The results are illustrated in 2. 4. These variations in case styles and boldness show a higher preference for the bold size in comparison to normal (not bold) for all forms of case style. The general ratio is about 60% bold size preference to 40% normal size preference and this result remains valid for a sans serif font (Arial) and a serif font (Times New Roman). The results also show that the case style has a slight deference regarding both used font (Arial, TNR).

Table 2.4: Boldness effect on text preference

Font	Size	All letters are lowercase	First letter is uppercase	All letters are uppercase
Arial	bold	49	42	49
	normal	31	38	31
Times New Roman	bold	45	45	47
	normal	35	35	33

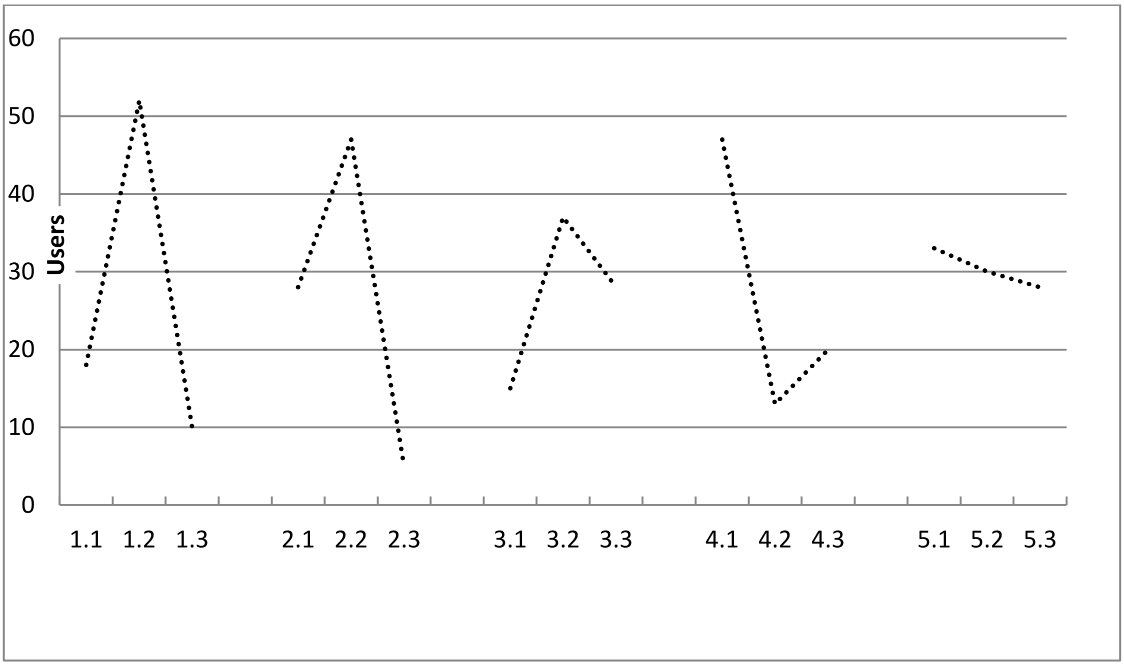
On the third level of complexity, size was obtained by varying case styles combined with other visual variables to present multiple levels of hierarchy (two or three levels). Complexity overlaps with hierarchy when different levels of hierarchy need to be presented. For such a level of complexity, three sequential font sizes correspond to three levels of importance. Here the highest level was completely uppercase, the middle level was designed with first letter uppercase and the lowest was fully written in lowercase letters (= 3 levels). This was later compared with a map containing labels with only two levels of hierarchy (= 2 levels); as the highest level was fully designed in uppercase and the other two were both labelled with first letter uppercase. These results are presented in Table 2.5. The simple hierarchy (2 levels) shows 20% more preferences than the one with 3 levels for an Arial font and 40% more preference for Times New Roman. By comparing the results of the levels of complexity and the difference between Arial and Times New Roman, the users’ preference shows the importance of simplicity when designing labels.

Table 2.5: The third level of complexity preference (with 2 and 3 levels of hierarchy)

Hierarchy	Arial		Times New Roman	
	2 levels	**3 levels**	**2 levels**	**3 levels**
Bold	48	32	56	24
Normal (not bold)	44	36	56	24

2.3.1.2 Shape

Shape preference was evaluated based on the frequency that participants chose for each studied shape. Different fonts were used to test shape preference. Since the number of shapes (fonts) is theoretically infinite, 15 different shapes were defined and divided into 5 groups (related to formality, spacing, serif, serif and spacing, sans serif and spacing) each with three levels. Each group thus tests one characteristic of shape which ranged from a low over middle to a high level. Figure 2.3 illustrates the user preference of each characteristic. This shows that users prefer fonts that have the moderate degree of formality, spacing and serif and prefer the lowest degree of serif and spacing and sans serif and spacing. When presenting fonts characteristics in a ranked order, it should be noted that the lowest level is preferred when two criteria of font are presented and the middle characteristic when individual criterion is presented.



fantasy	*1.1*	spacing 0	*2.1*	serif 0	*3.1*	serif & spacing 0	*4.1*	sans serif & spacing 0	*5.1*
formal	*1.2*	spacing 1	*2.2*	serif 1	*3.2*	serif & spacing 1	*4.2*	sans serif & spacing 1	*5.2*
historical	*1.3*	spacing 2	*2.3*	serif 2	*3.3*	serif & spacing 2	*4.3*	sans serif & spacing 2	*5.3*

Figure 2.3: Shape preference of the same texture

Three methods of labelling according to case style were tested for the two most frequently used types (Times New Roman and Arial). Upper case, lower case and first letter upper case labelled map were tested pair wise. The arguable theory of serif and sans serif use of typography was thus tested. Table 2.6 represents the users' shape preference showing the highest preference for Arial with about 70-80% versus a 20-30% preference for Times New Roman.

Table 2.6: The influence of shape on text preference

Case style	Times New	Arial
All letters lower	20	60
First letters upper	27	53
All letters upper	15	65

2.3.1.3 Orientation

The orientation of typographic symbols was studied in two phases. First, the orientation of the entire word was evaluated by the participant. The orientation of the labels used during the study corresponds to a typical output of most cartographical software: all horizontal, all tilted (under the angle of the largest diagonal) or mixed (horizontal if applicable within the boundaries, otherwise tilted). The participants tended to prefer the horizontal orientation, followed by the tilted orientation and finally the mixed type orientation. This is listed in Table 2.7. Under the circumstances of the experiment, it was found that the users' preference of horizontal matches the orientation of text reading.

Table 2.7: Orientation influence on text preference

Orientation	Participants	%
Horizontal	59	74
Tilted	10	12
Mixed	11	14

In the second phase, the orientation of each letter was considered. The italic style was tested with the three levels of complexity (see Table 2.8). At the first level of complexity, straight and italic was tested for four font sizes of the labels (8, 10, 12, 14). The straight orientation was preferred over all the tested sizes (8, 10, 12, 14). The second level of complexity involved italic versus bold italic in pair wise comparison. In this case, only the smallest size acquired the highest preference for bold italic, while the opposite is true for italic (with also a rather high value for medium sizes). For the third level of complexity, the result is exactly the opposite. Bold italic was tested versus narrow bold italic in this case. The smallest font 8 has the highest preference for bold italic whereas the preference of narrow bold italic is for the larger font sizes 10, 12, and 14. It is important to note that the users' preference of second and third level of complexity is thus dependent on font size.

Table 2.8: Complexity effect on orientation of text preference

Font size	First level of complexity		Second level of complexity		Third level of complexity	
	Straight	**Italic**	**Italic**	**Bold italic**	**Bold italic**	**Narrow bold**
8	65	15	17	63	64	16
10	42	38	47	33	38	42
12	47	33	47	33	17	63
14	74	6	70	10	9	71

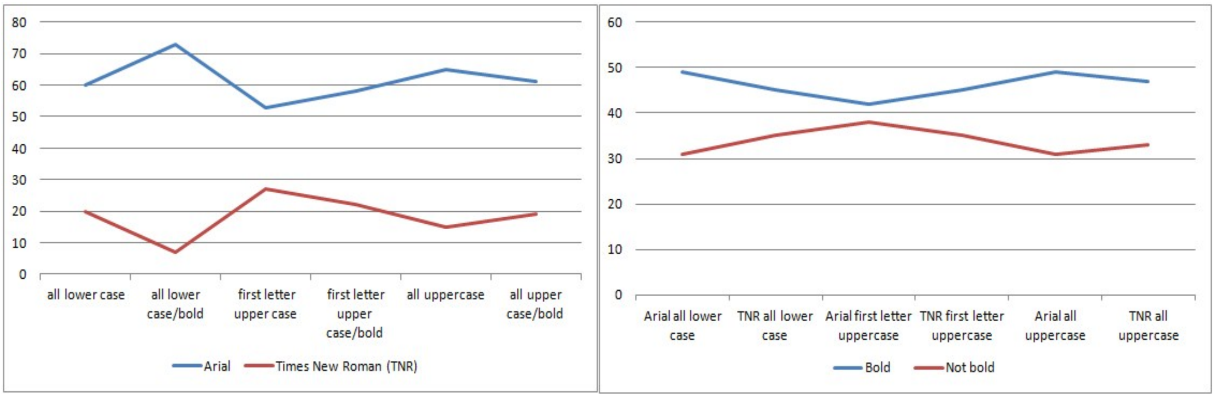


Figure 2.4: The preference of shape and size of different textures

2.3.1.4 Texture

The design of texture has a wide array of choices as they are based on a combination of visual variables. As a consequence, the textures were divided according to font type: serif (Times New Roman) and sans serif (Arial). These two main groups were further subdivided according to the applied variables: textures of the same font (shape) and same size, multi-texture of the same font and different sizes and finally multi-texture of multi-font and different sizes. For each mentioned division the users identified which one they preferred, which is shown in Figure 2.4. Using one way ANOVA, the test declared that no statistical significant deference was detected for the three groups of texture and between the two font types (F=0.00, P=1.00). However, when only one font was considered, the texture designed of Arial was more preferred. In addition to that, the textures designed of bold fonts were more preferred.

2.3.2 Experience and gender influence

Both clusters of experience and gender were tested by asking participants to show their preference of 79 sets of two or three map combinations. These sets are devoted to test size, shape, orientation and texture. For each set users' preference was analysed statistically using two-sided T-test.

2.3.2.1 Experts versus novices

Expert participants are well trained to use maps and work with maps on a daily basis. It was expected that their attitude towards map symbols, especially the typography, would therefore differ from the novice users. The ranking test was presented to both user groups (experts and novices). Shown in Table 2.9, using two- sided T-test only three statistical significant different were found. They are emerged in the variables of orientation, shape and texture.

Table 2.9: The detected level of preference variation between experts and novices

CRITERIA (Experts/ Novices)	**Pearson Chi-Square Value**	**Asymp. Sig. (2-sided)**
Arial-narrow-10 size-bold vs. Arial-narrow-10 size-bold-italic	3.63	0.049
Arial-bold-all uppercase vs. TNR-bold-all uppercase	7.735	0.005
Arial-normal-2 levels vs. TNR-normal-2 levels	4.832	0.028

Figure 2.5 illustrates these significant differences between the experts' and novices' preference. The variation between the two groups occurred in the sub-case of orientation Arial narrow bold italic 10 size. The test showed a fair significant difference (0.049). This draws attention to the care that should be taken in labelling with 10 size as the most preferable and readable size. The second significant difference was the same texture (all letters were in uppercases), although with a different font. A high significant difference was recorded (0.005) as the experts preferred Arial font type for this task, whereas the novices preferred Times New Roman (TNR). This difference between a serif and a sans serif font must be taken into account when creating maps for different user groups. Similarly, the same texture with two levels of hierarchy and different bold fonts showed a high significant difference (0.028), with a majority preferring Arial.

2.3.2.2 Females versus males

Concerning the preferences of females and males, data were compared pair wise between the two user groups. Four significant differences were recorded when using T-test (see Table 2. 10). Starting from size 10 bold versus 14 bold, a high significant deference was located (0.025) as females rejected the larger size and preferred the 10 size. This case was recorded only for the bold size, which in turn demands much more attention and care for the association with other variables. The second significant difference was located between Arial narrow 12 bold and Arial narrow 12 bold italic (0.042). The third significant difference (0.047) was detected for the uppercase which enlarges the symbols and makes the details more visible, in addition to the difference between Arial and Times New Roman

(TNR). The final significant difference regards the same texture and different font of Arial and Times New Roman was highly significant (0.005). The texture used for this test was a combination of the three levels of hierarchy and boldness for size. The result was a decreased preference for Times New Roman by female participants. See Figure 2.6.

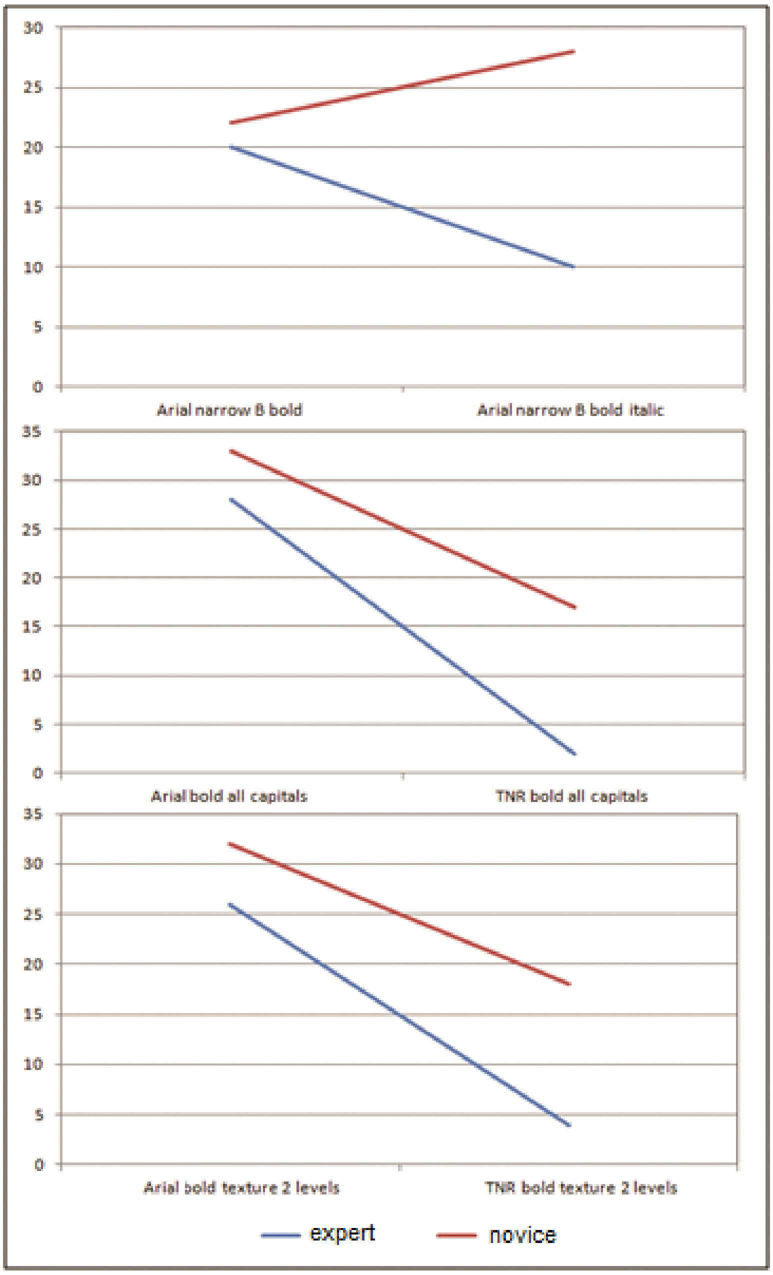


Figure 2.5: Significant differences between experts and novices

Table 2.10: The detected significance levels for gender preference variations

CRITERIA (Females/Males)	Pearson Chi-Square Value	Asymp. Sig. (2-sided)
Arial-14 size-bold vs. Arial-10 size-bold	5.045	0.025
Arial-narrow-12 size-bold vs. Arial-narrow-12 size-bold-italic	4.127	0.042
Arial-all uppercase vs. TNR- all uppercase	3.940	0.047
Arial-bold- 3 levels vs. TNR-bold-3 levels	8.061	0.005

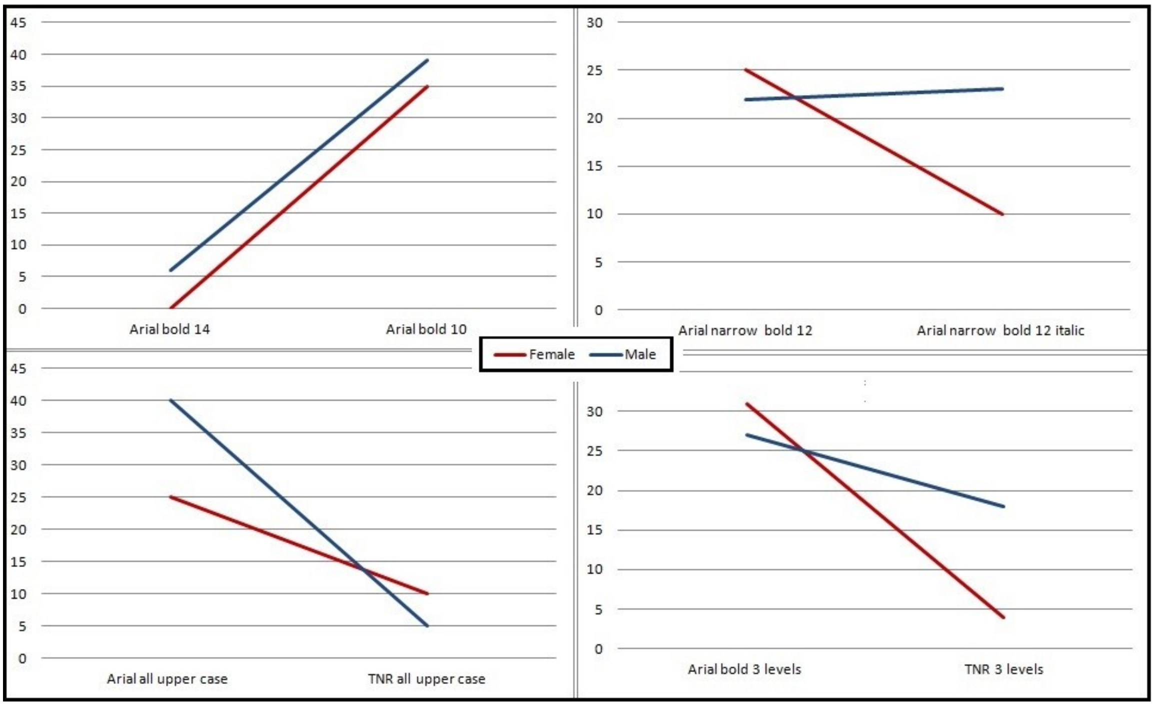


Figure 2.6: Located significant differences for females and males

2.4 Conclusions and Future perspectives

It is not possible to define one typographic variable as the 'best' variable for all typographic functions and for all map users. Although it has been found that some criteria have much more acceptance than others, it is still unclear which is the most efficient and legible. To acquire these insights, further experiments are required with time measurement as a scale of efficiency.

The levels of complexity utilised declared few remarkable differences for some criteria, such as narrow typography, italic and bold. Additionally, contradictions were found when using complex textures compared to font size. This finding is due to the fact that boldness makes the symbol larger and narrower. It has also been concluded that when size is combined to other variables such as narrow and italic the larger size is highly preferred over the smaller.

For the three levels of complexity, different typographic criteria are the most preferable and legible for the examined group. Bertin's preferred variable on typography can be classified as follows:

According to the map scale used and the distribution of labels illustrated previously in Figures 2.1 and 2.2 (including point data and area data), font size 10 is most preferred by participants working with screen maps.

A bold font allowed the typographic symbols to stand out more, thus making it more eye-catching for the participant. A higher level of preference for bold fonts, as illustrated in Figure 2.4, confirms this finding. In maps of a higher complexity, however, bold fonts tended to be a detrimental variable.

Arial is the most preferred font type as an individual variable. This finding could limit some cartographic design. However, this result is not only true when the font is varied individually; but also when font (shape) is associated with other variables such as size and texture, Arial type is generally preferred.

It is remarkable that the preference of different shapes followed almost the same trend in the analyses of both experience and gender. Text simplicity or complexity represented by font showed no specific significant difference, meaning that the preference trend for each group could be explained by the overall trend of shape preference and vice versa.

Horizontal orientation is the most preferable typographical variable. However, it has the disadvantage of assigning the orientation of the geographical features. Additionally, the text might exceed the available space for a small feature size. According to the cartographical rules and needs, map designers cannot always utilise horizontal orientation.

The orientation of the word letters (straight versus italic) varies depending on the level of complexity design. For the first level of complexity, users show noticeable preference of the straight orientation mean while it was remarkably noted that for the second and third level of complexity the user preference depend on the labels font size.

For point data labelling, it is critical to establish rules as it has been shown that differing groups have varying tendencies, especially when it comes to the second and third levels of complexity.

The results indicated some statistical difference within the experience and gender clusters. What is remarkable is that these differences were not about variables of the first level of complexity but about the second and third one.

The visual variables of order perception (size, value and texture) can be used individually or combined to introduce hierarchy. This research presented hierarchy by ordered variation in size (font size, boldness and case style). Using two levels of hierarchy was remarkably preferred over using three levels of hierarchy. This result remains valid for Arial (sans serif) and TNR (serif) fonts.

The preferences of the examined groups unsystematically vary between them and according to the studied variable, which requires more standardising techniques and other studies to acquire more detailed insights in the legibility and efficiency of typographic symbols. Additionally, the results are only valid on a blank background where the distribution of features is the most important function of the map and thus the primary function of text is to provide the geospatial address. Obtaining more empirical information on the relationship between labels and the surrounding map elements is essential and will be included in further studies, such as different background colours.

However, care should be taken when studying colour and value. They are very critical variable because of the interaction between the colour of labels and the colours of map elements in first place. Secondly, the medium on which the map will be presented is critical as well. Therefore, a thorough study is planned to obtain insights in the relation between label visualisation and the map background (topographic, thematic and their use of colours): which value is the best for a certain layout, the influence of colour onto map legibility and others.

Determining the best legible typographic variable is a combination of two tasks, starting with map aesthesis and ending with map efficiency, including a series of rules and priorities regarding map design. Further studies need to be undertaken in order to explain the relation between text functions and their design, which plays a critical role for utilising different variables individually or combined.

Furthermore, it would be interesting to extend this work to the applications of label on other language symbols than Latin in order to test the influence of visual variable on different lettering systems such as Chinese or Arabic. Moreover, the efficiency of visual variable on typography will be tackled in future research.

2.5 References

Bartz, B.S. 1970b. Experimental Use of The Search Task in The Analysis of Type Legibility in Cartography, The Cartographic Journal, 7(2), pp. 103-113.

Bartz, B.S. 1970a. An Analysis of the Typographic Legibility Literature, Assessment of Its Applicability to Cartography, The Cartographic Journal, 7(1), pp.6-17.

Bertin, J. 1967. Semiology of Graphics: Diagram, Networks, Maps, Translated by William J. Berg, 2011, Esri press, California.

Board, C. & Taylor, R. M. 1976. Perception and Maps: Human Factors in Map Design and Interpretation, Transactions of the Institute of British Geographers, 2(1), pp.19-36.

Fairbairn, D. J. 1993. On The Nature of Cartographic Text, The Cartographic Journal, 30(2), pp.104-111.

Foster, J. J & Kirkland, W. 1971. Experimental studies of map typography, Bulletin of the society of university cartographers. 6(1), pp.40-45.

Gerber, R. 1981. Competence and performance in cartographic language, The Cartographic Journal. 18 (2), pp.104-111.

Imhof, E. 1975. Positioning Names on Maps, The American Cartographer, 2(2), pp.128-144.

Koch, W. G. 2001. Jacques Bertin's Theory of Graphics and Its Development and Influence on Multimedia Cartography, Information Design Journal, 10(1), pp.37-43.

Kraak, M. J. & Ormeling, F. 2010 Cartography. Visualization of Spatial Data, 3rd edition, Person Prentic Hall, Harlow.

Kraak, M. J., Edsall, R. & MacEachren, A. M. 1997. Cartographic Animation and Legends for Temporal Maps: Exploration and or Interaction, Proceedings 18th ICA/ACI International Cartographic Conference, Stockholm, Sweden.

MacEachren, A. M &Taylor, D. R. F. 1994. Visualization in Modern Cartography, Pergamon, Oxford.

MacEachren, A.M. 1995. How Maps Work, Representation, Visualization and Design, the Guilford press, NY.

Phillips, R. J., Noyes, L. & Audley, R. J. 1977. The Legibility of Type on Maps, Ergonomics, 20(6), pp.671-682.

Robinson, A. H. & Petchenik, B. B. 1976. The Nature of Maps, University of Chicago press, Chicago.

Slocum, T. R., McMaster, R. B., Kessler, E. C., & Haward, H. H. 2005. Thematic Cartography and Geographic Visualization, 2nd edition, Person Prentic Hall, Upper Saddle River, NJ.

3

Toward a Deeper Understanding of Cartographic Text Visualisation: Assessment of User Preferences and Colour Influence

Modified from: Deeb, R., Ooms, K., Van Etvelde, V. & De Maeyer, P., In press. Toward a deeper understanding of cartographic text visualisation: assessment of user preferences and colour influence. The Cartographic Journal. DOI: http://dx.doi.org/10.1179/1743277413Y.0000000073

Preface

The main goal of this chapter is to investigate the influence of a map's background colour on user preferences regarding labels' typography, including their size, shape, orientation, and texture. Four sets of backgrounds were tested: blank maps, grey scale, hot colours, and cold colours. The foreground of these maps is populated with name labels, which are associated with point objects or areal objects. Bertin's visual variables were applied to this set of name labels, both separately and in combination. User preferences of different typographic variations were registered and compared on the basis of different background colours using a one-way ANOVA. The results indicated that the typographic design of the map labels should not be adapted according to the map's background colour.

3.1 Introduction

A map is a composition of feature representations, which are generalised and visualised using three object types: points, lines, and areas. Each of these objects is symbolised and visualised using a set of visual variables: colour, orientation, size, shape, texture, and value. Bertin (1967) studied these visual variables and linked them with perceptual properties (associativity, selectivity, order, and quantity) to set rules for their appropriate use. The goal of Bertin's rules was to improve the design of the individual elements on a map so that map users could interpret and understand the map's content more efficiently. Combinations of these visual variables are frequently used within one map element; for example, a thicker blue line indicates a more important (size indicating order) river (colour indicating selection). Over the years, extensions to these basic rules have been proposed, especially with the rise of three-dimensional and dynamic representations (MacEachren, 1995). Garlandini and Fabrikant (2009) rightly noted that very few of these long-standing cartographic design principles had ever been tested with actual map users. They concluded from their experiment that the visual variable size was considered the most efficient and effective variable for visual communication.

In addition to points, lines, and areas, name labels are an object type that is often overlooked. These labels can present much more information than other types of symbols, as they contain two 'levels' of information. First, they inform the map reader about the toponym of a certain geographical location, information that cannot be presented using another type of symbol. Second, their typography and placement may stress the hierarchy, classification, and/or spatial structure of the associated object (Imhof, 1975). Fairbairn (1993) considered text to be an indispensable (fourth) map element in addition to points, lines, and areas. He extensively described the different purposes that text may have on maps. Normally, text is placed on top of other map layers representing, for example, thematic information such as property values. As a consequence, these underlying map elements -and especially their colour- may influence the user's perception of the text.

Bartz (1970) discussed the influence of typographic size and its appropriate dimensions using time measurements. Her study was limited to topographic maps printed on paper. However, she concluded that the figure-ground relation, location on the map, relation and contrast with other labels are much more important than size variations when locating target labels. Starting from Bertin's work on visual variables, Kraak and Ormeling (2010) suggested a set of rules related to the correct use of typography on maps. These rules concern hierarchical and nominal labelling presentation; the authors proposed boldness, size, spacing, colour value, and case style to indicate hierarchy. In addition, they proposed using colour, shape, and italicisation to indicate nominal differences. The influence of word crowding, which is affected by line spacing and word spacing, on the speed of reading plain text was studied by Yu *et al.* (2011). In addition, demonstrating the legibility of serif and sans serif fonts was the main concern of Arditi and Cho (2005), who investigated normal reading tasks on plain text. The serif font was found slightly more legible than the sans serif font. Sans serif fonts were tested by Feldmann and Kreiter (2006) on topographical maps. Laboratory tests of Arial, Univers, and Frutiger revealed similar reading speeds for the sans serif fonts. The variation in the typographic preference of different users considering their experience and gender was studied by Deeb *et al.* (2012). The authors used blank maps and located some differences between the investigated user groups.

Based on the work of Bertin (1967), text characteristics -such as size, colour, spacing, upper and lower case- can be linked to a certain visual variable and thus to certain characteristics of perception by the map reader. Therefore, the function of a label is determined by its own design and influenced by the surrounding elements. The combined application of visual variables on map objects may cause unwanted influences on the interpretation of the objects themselves or on neighbouring objects. Special care should be taken when using colours. Cleveland and McGill (1983), for example, investigated the influence that colours have on users' perception of object sizes. They used a simple thematic map representing counties in the state of Nevada (USA). Different colours were used to create different stimuli. The authors concluded that the use of colours on maps can cause optical illusions: areas represented in red were thought to be larger despite having the same size.

The colour of the objects in a choropleth map may also influence the perceived colour of adjacent objects, due to lateral inhibition. When two objects are represented with the same shade of grey, they will appear differently if they are surrounded by objects with a darker shade or a lighter shade of grey (MachEachren, 1995; Monmonier, 1996). Brewer (1996) suggested avoiding the combination of complementary colours within one map. She discovered that if an object is surrounded by objects with a different colours, its colour will shift towards the complement of the surrounding colour. A number of user studies focused on determining users' preferences regarding the use of certain colours or colour combinations (Granger, 1955; Ou *et al.,* 2004 and others). Ou *et al.* (2004) studied the relation between the preference of colour combinations on the one hand and colour emotions or colour harmony on the other hand.

However, previous studies did not consider the relation with other map objects. Steiniger and Weiber (2007) investigated the horizontal relationship between objects and their relation to map constraints and cartometric measures including geometric, topological, semantic, statistical, and structural properties of map objects. They proposed a typology to formalise these relations, which can be used as an aid in the automated generalisation. This study, like most other related studies, did not consider text to be an independent object.

Most studies mentioned earlier investigated users' performance measures, not their preference of typography given a certain stimulus (map in this case). Wood (1993) described different levels at which a users can express their preference. He considered preferences as 'aesthetic' responses. The first level is non-functional and purely artistic. In this case, the map is seen only as a piece of art. At the second level, the overall structure and balance of the map and its composing elements are considered. These elements include symbolisation, colour, and typography (Karssen, 1980; Wood, 1993).

This chapter considers the users' preference response at both levels. The users' preferences towards a certain functional typography of labels are studied in combination with varying background colours of the maps. Thanks to modern cartography and display methods, maps can be visualised and manipulated on screen. Because the old typographic research considered only paper maps, it is important to consider new aspects that are inherently related to digital maps, such as (limited) screen sizes and resolutions (2003). When considering the quality of a website (which may include maps), user preferences cannot be ignored. The number of websites that currently contain maps is unmeasured and increases every day. Therefore, preference is an important aspect when designing maps for websites to ensure that users continue to visiting them. The goal of studying user preferences (regarding the design of texts) is the improvement of the maps' design to facilitate the communication process: a certain message must be transferred to the map reader. The main goal of this chapter can be subdivided into two objectives. First, it is important to identify the impact of using different options regarding visual variables on the map users' preference. Second, the study investigates the influence of the map background onto users' preference for the typographic foreground. In the next sections, the design and results of this study are described in detail.

3.2 Method

An experiment was conducted to examine the influence of map background colours on users' preference of the typographic design. This user study utilised choropleth and topographic maps as backgrounds and different label typographies on the foreground. This combination was made for four background colour designs.

3.2.1 Participants

Four groups were formed, each containing 50 participants. To maintain homogeneous characteristics along with the four groups and prevent any bias due to participants' characteristics (Nielson, 1993; Rubin and Chisnell, 2008; and Aykin, 1989), all groups had 25 participants with a high level of experience (experts) and 25 participants with a low level of experience (novices). The expert users had obtained at least a master degree in Geography or Geomatics or had a high interest in cartography, its applications and research fields. The novice users were students who had just begun their BSc education. None of them received any cartographic training prior to the study. In addition, it was attempted to keep the gender balanced: there were 25 females and 25 males in the first group, 26 females and 24 males in the second group, 23 females and 27 males in the third, and 23 females and 27 males in the fourth group. Furthermore, all participants had normal vision or vision that had been corrected to normal.

3.2.2 Stimuli

In total, 41 different maps were presented to each participant, using a questionnaire presented on a screen (Lime Survey package was used). Each of these digital maps was combined with one or two other maps so that the participant could compare them, resulting in a total of 66 different questions. Four different questionnaires were made, related to the background colour of the maps. Map combinations always had the same type of background but differed in the label typography (size, boldness, font, etc.). Four phases of the experiment were made with the four described map backgrounds. The four different types of maps are illustrated in Figure 3.1 and designed as follows:

- Maps whose background is blank (Blank Map, BM): only point objects or area boundaries were depicted in the background. As a consequence, this type of map will not experience any influence from background colours (Hue=0°, Saturation=0%, Value=100%; See Figure 3.1, top row).
- Maps whose backgrounds are depicted in different shades of grey (Grey Scale, GS), illustrated in Figure 3.1, second row.
- Maps whose backgrounds are filled with colours whose hue ranges from 0°-90° and 270°-360°. Based on the colours contained in these intervals, these will be called Hot Colour (HC). Figure 3.1, third row illustrates examples of these maps.
- Maps whose backgrounds are filled with colours whose hue ranges from 90°- 270°. Based on the colours contained in this interval, these will be called Cold Colour (CC). Examples are illustrated in Figure 3.1, last row.

To avoid measurement biases due to resolution and size differences, all participants completed their test on a flat screen with 1280x1024 resolution and a size of 17 inches. The stimuli were presented in the same order to the participants of the four questionnaires. The visual variables were grouped together in individual blocks. Therefore, the users can determine their preference between different options of the same variable (for example, different sizes). The design of the maps will be discussed in details in the next sections.

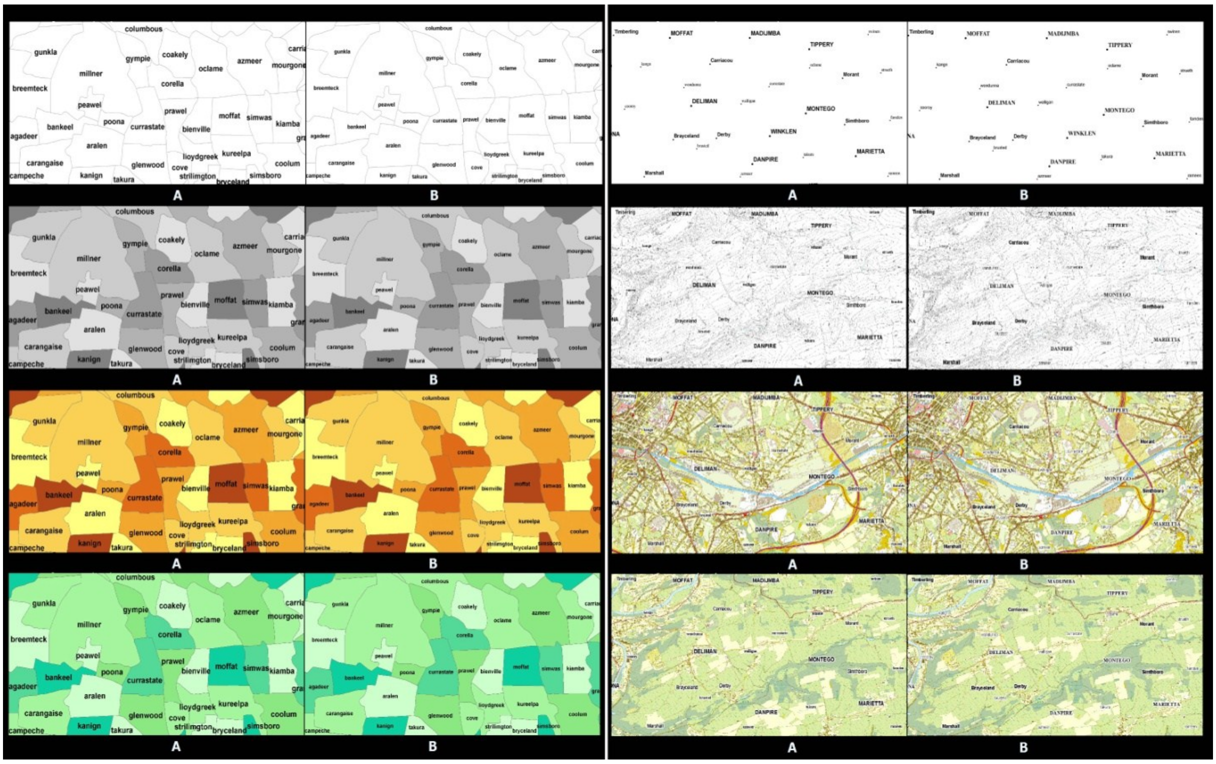


Figure 3.1: A comparative example of test stimuli of the four background types, Areal labelling (left) and point labelling (right)

3.2.3 Map design

Labels' designs on the maps (stimuli) varied in size, shape, orientation, and texture. An extensive explanation of these visual variables can be found in the work of Bertin (1967), and their application on map labels was described by Deeb *et al.* (2012). These different design criteria were applied individually or combined (for example, a serif font and a large corpus). Fictive toponyms were placed on the map to ensure that participants would not recognise the location, as their ability to do so could bias the results due to previous knowledge. The position and content of the labels were equal on all maps, but their length varied between the maps. Two different labelling options were considered: labels associated with areal objects and labels associated with point objects. To visualise different categories of the labels and their associated (point or areal) objects, the graphical variables of size, shape, and texture were applied to the associated name labels. In addition, variations in the labels' orientation were also tested on choropleth maps.

As mentioned earlier, four different groups were formed based on the background colour of the maps: blank maps (BM), grey scale maps (GS), hot-coloured maps (HC), and cold-coloured maps (CC). Demonstrated in Figure 3.1, the maps with point data were based on the Belgian topographic map series of 1: 10 000. The hot-coloured maps, on the one hand, show urban areas which were mainly depicted in colour hues ranged between [0°-90°, 270°-0°], including yellow, orange, and red. The cold-coloured maps, on the other hand, show rural backgrounds whose colour hues ranged between [90°-270°], including shades of blue and green.See Table 3.1.

Table 3.1: HSV parameters of the map designs

COLOUR	1st class			2nd class			3rd class			4th class			5th class		
	GS	HC	CC	GS	HC	CC	GS	HC	CC	GS	HC	CC	GS	HC	CC
Hue °	0	60	119	0	45	115	0	37	115	0	25	150	0	18	156
Saturation%	0	49	20	0	66	36	0	81	44	0	89	61	0	86	93
Value%	90	100	100	80	98	96	70	95	91	60	88	85	50	70	81

The choropleth maps (depicting areal data) were visualised by five classes, ranging from a light to a darker colour, illustrated in Figure 3.1, left column. The hot-coloured backgrounds thus corresponded to an orange colour ramp; the cold-coloured maps corresponded to a green colour ramp. The blank maps with areal data only showed the boundaries of the areas without any differentiation of the five classes.

3.2.4 Task

The stimuli were implemented in an online questionnaire that allowed users to express their preference for a certain map (design). In each question, two or three maps were presented simultaneously, and participants could rank these maps according to their preference. On average, this task was completed within 20-30 minutes. The result was an ordered list of maps for each participant, indicating a personal preference for the application of graphical variables on map labels. Four sets of data were collected, each related to one of the background types (blank, grey, hot colour, and cold colour). This design enabled a comparison of the preference data of the four background types.

3.3 Results

In the results, the users' preference for certain label typography between the four map background colours was compared. In addition, a distinction was made between labels associated with areal data and those associated with point objects. The results of the user study are described in the following section, structured according to object type (point or area) and typography.

3.3.1 Labels associated with areal data

As described before (see Section 2.3), the areal data were presented by choropleth maps, and four different colour ramps were used: no colour ramp (BM), grey colour ramp (GS), orange colour ramp (HC), and green colour ramp (CC). The graphic variables that were applied to labels associated with areas were as follows: size, shape, texture, and orientation. Size can be applied to texts by using a different corpus, using bold texts, and using case styles. The shape of a label was determined by its font style, font family (serif versus sans serif), letter spacing, or a combination of these. Variations in texture were obtained by combining typographic options such as bold, narrow, and italic. Orientations can be studied within the labels (italic versus straight) or by varying the angle of the labels. A more detailed description of the application of these variables on the labels' design can be found in a study by Deeb *et al.* (2012).

3.3.1.1 Size of the labels

The test name labels were designed in Arial using four levels of dot sizes: 8, 10, 12, and 14. A bold Arial font was added, as it also enlarges the characters of the labels. These sizes were compared pair-wise for the four types of background. The frequency of users' preference of each size was counted and analysed statistically.

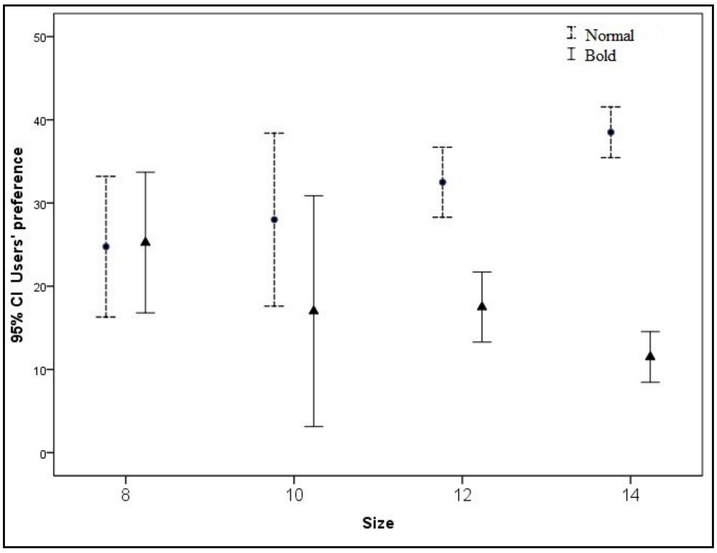


Figure 3.2: The mean users' preferences on both size groups (normal & bold) regarding the four dot sizes

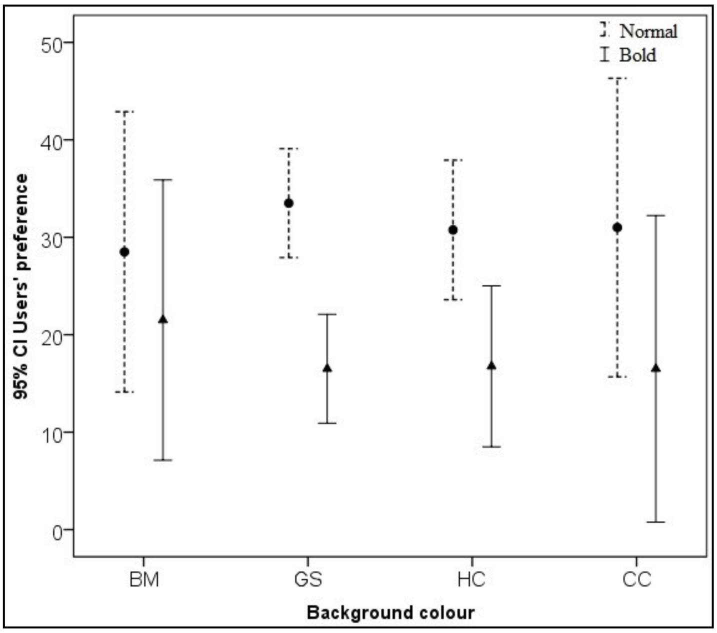


Figure 3.3: The mean values of users' preferences on size (normal and bold) regarding the maps background colours

The study of the font sizes was made in two steps. First, the difference between bold size and normal (within a same dot size) was considered, and users' preference was compared between four corpus sizes (both bold and normal). Because boldness enlarges the size of the label, bold labels were compared pair-wise to normal labels separately for each of the four map backgrounds. The results, depicted in Figure 3.2, show the general trend in user preferences for the different label sizes for all background colours. Figure 3.2 illustrates the mean user preferences for each dot size. This graph shows that the users did not prefer a large label size (12 or 14 dots) in bold as much as they did for normal size labels. The preferences regarding the smaller font sizes (8 and 10 dots) are less different.

These results of the four background maps were tested pair-wise (normal versus bold) for each individual dot size using one-way ANOVA; User preferences did not differ significantly between size 8 normal and 8 bold (F=0.018, P=0.899), and user preferences did not differ significantly between size 10 normal and 10 bold (F=4.079, P=0.090). Meanwhile, user preferences for size 12 normal and 12 bold were significantly different (F=64.286, P=0.000). In addition, user preferences for size 14 differed significantly between normal and bold (F=394,636, P=0.000). Overall, the users appeared to have a higher preference for normal typography. Using a one-way ANOVA, a significant difference was found across different sizes for normal fonts (F=13.006, P=0.000) and bold fonts (F=11.147, P=0.001). In addition, a one-way ANOVA test was conducted for both normal sizes and bold sizes considering the background colours, and no significant difference was found for the former (F=0.125, P=0.943) and (F=0.258, P=0.854) for the later (see Figure 3.3). In view of the pair-wise comparison of normal and bold sizes of the four background colour, two significant differences were located for grey scale maps and hot-colour maps when using one-way ANOVA, (F=64.865, P=0.000) and (F=16.622, P=0.007), respectively. Meanwhile, no significant differences were found for normal and bold blank maps (F=1.200, P=0.315) and cold-colour maps (F=4.419, P=0.08).

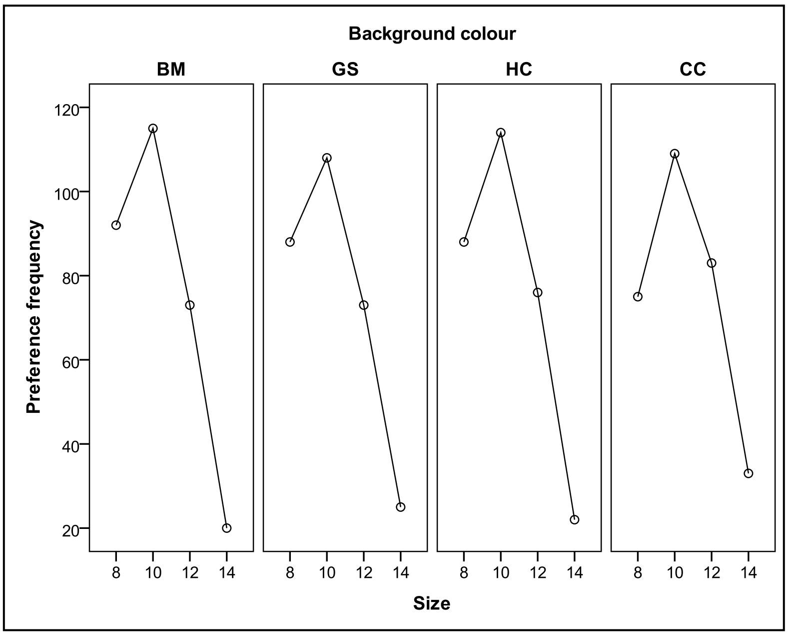


Figure 3.4: Trends of the frequency of users' preferences of labels' size regarding to background colour

Second, the user preference of the four dot sizes (8, 10, 12, and 14) was compared pair-wise. The preference regarding the size was analysed and compared along the four map backgrounds. The frequency of each size preference was counted and accumulated to have a weighted value of size preference regarding the four background types, which is illustrated in Figure 3.4. The graph shows the same peak in the preference data for all background colours: a size of 10 dots has the heights preference. Approximately 55% of the users preferred this size in the four map types (blank map, grey map, hot-coloured map, cold-coloured map). These percentages are calculated based on a maximum frequency of 200 (50 users, 4 colours) for each size. The lowest preference (approximately 10%) was associated with size 14 dots. The accumulated data were statistically analysed using a one-way ANOVA. These tests showed no significant difference in users' responses between the four sets of maps (F=0.11, P=0.95).

3.3.1.2 Shape of the labels

To acquire a clear overview of the relationship between the map background colour and the preference in the labels' shape, a fixed corpus size was used in combination with different typefaces. These variations in shape were obtained using a combination of the following parameters; formality (artistically different ranging from the old hand writing to machine fonts), (sans) serif, and spacing. For each shape group, three options were depicted simultaneously to the user, who could rank the three maps according to one's preference. The exact fonts that were used (each with a size of 12 dots) in each group are:

- Formality: (A) Blackadder – (B) Comic Sans SM – (C) Times New Roman;
- Spacing: (A) Arial Bold – (B) Times New Roman – (C) Sea Black;
- Serif: (A) Comic Sans SM – (B) Times New Roman – (C) Century Schoolbook;
- Spacing & serif: (A) You Yaou – (B) Times New Roman – (C) Century Schoolbook;
- Spacing & sans serif: (A) Arial Narrow Bold – (B) Arial Bold – (C) Comic Sans SM.

The users' preferences for each of these options and for the four backgrounds are depicted in Figure 3.5. However, users had a distinct preference for a certain label shape among the three options shown simultaneously. A one-way ANOVA test was used to investigate users' preferences towards the three categories of each group where some significant differences of their choices were located. The preference of formality categories was significantly different ($F1=45.079$, $P1=0.000$), as users preferred Times New Roman the most. In addition, the degree of preference within serif categories was significantly different ($F3=49.4050$, $P3=0.000$). Noticeably, users preferred Times New Roman among this group's options. Furthermore, the preference of spacing & serif categories was significantly different ($F4=86.470$, $P4=0.000$) with a significant preference of You Yaou. Moreover, the preference of spacing & sans serif categories was significantly different ($F5=8.908$, $P5=0.007$), as Arial Bold was significantly the least preferred shape. On the contrary of that, users' preference of spacing categories was not significantly different ($F2=3.915$, $P2=0.06$).

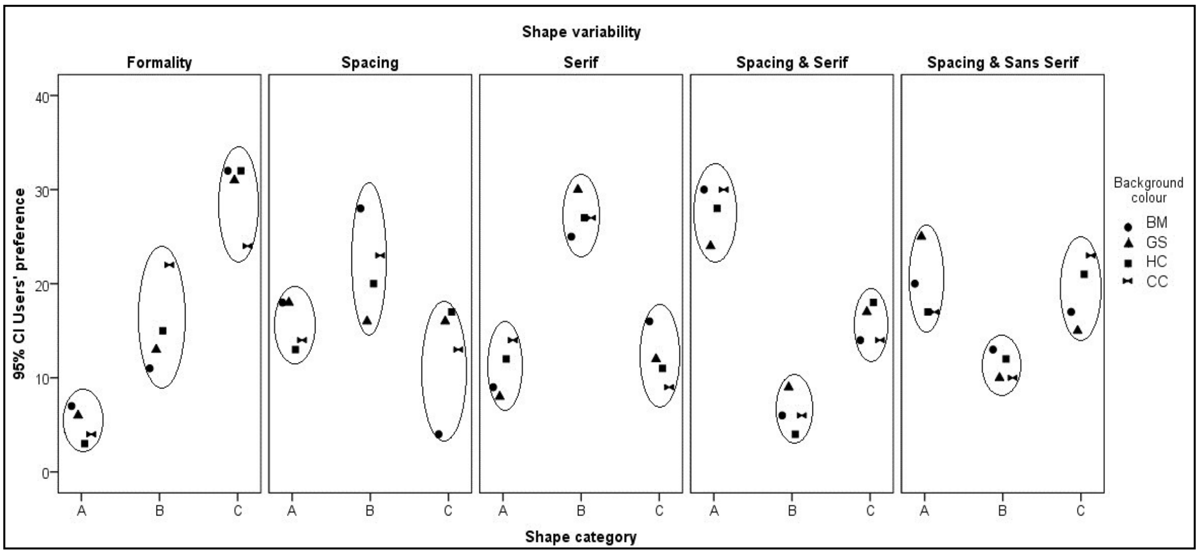


Figure 3.5: Users' preference on shape categories regarding background colour

Illustrated in Figure 3.5, the preference values were compared statistically across the four background types for each of the group options. Using a one-way ANOVA, no significant difference was found regarding the background colours for the five proposed options; formality, spacing, serif, spacing & serif, and spacing & sans serif (F1=F2=F3=F4=F5=0.000, P1=P2=P3=P4=P5=1.000).

3.3.1.3 Texture of the labels

Texture can be seen as a set of consistent graphic patterns. The texture of labels was varied using a combination of three parameters: italic, bold, and narrow. Three combinations were applied to the labels using an Arial font with four sizes (8, 10, 12, and 14 dots). The participants had to express their preference for each of these combinations: italic versus bold italic, bold versus narrow bold, and narrow bold versus narrow bold italic.

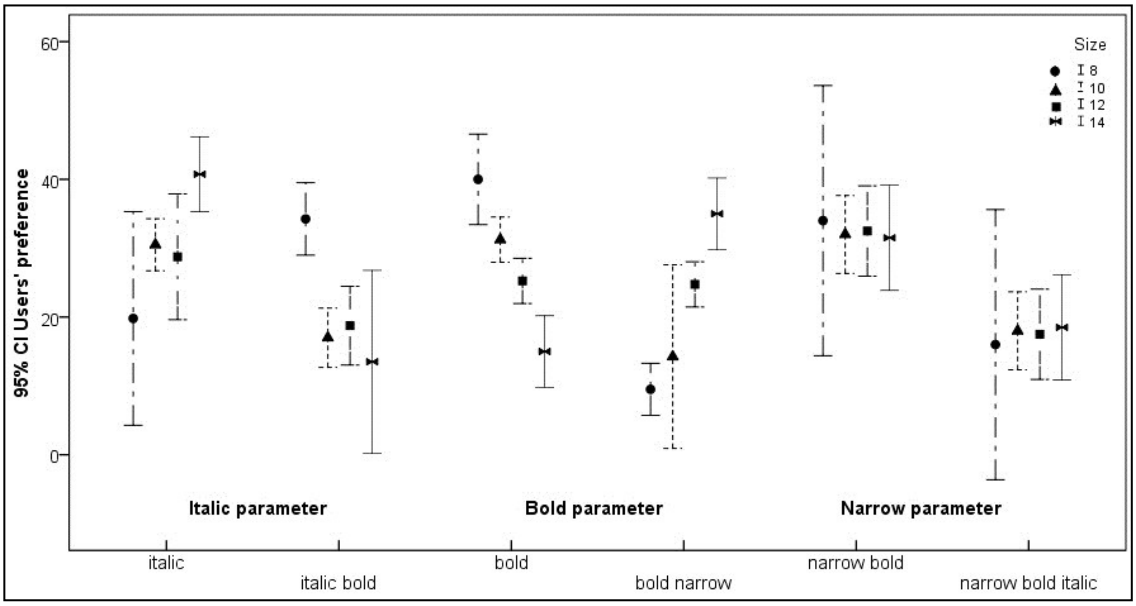


Figure 3.6: The mean values of users' preferences of texture variable

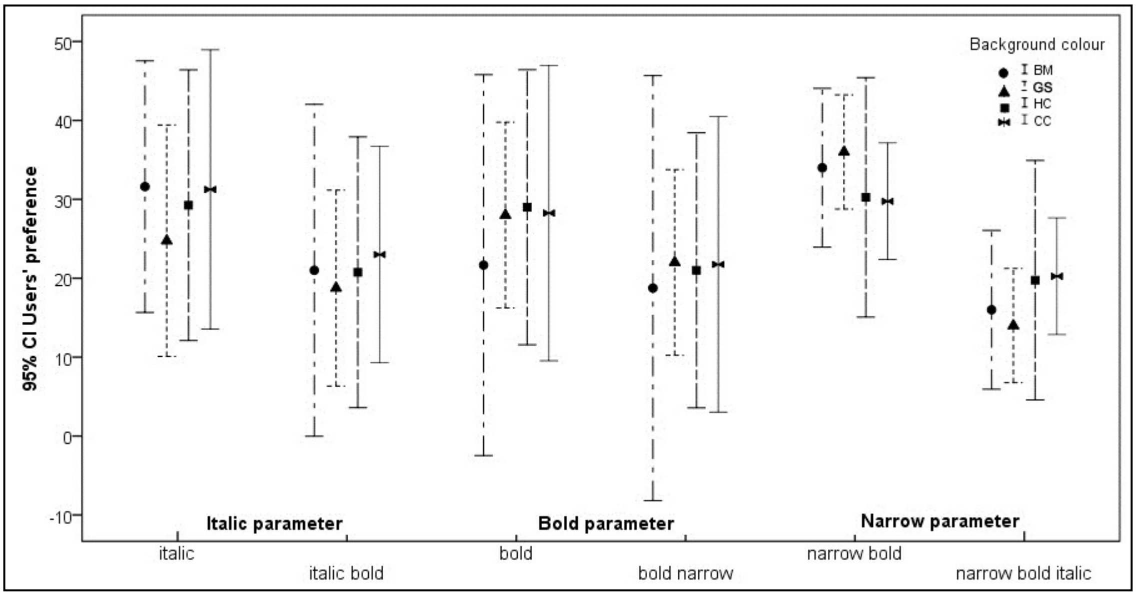


Figure 3.7: The influence of background colour onto the mean preference of texture variable

Figure 3.6 illustrates the mean values of user preferences regarding the labels' texture. An ANOVA test was used to compare the variation of the textures described before where no significant difference was found for the first comparison (F1=3.724, P1=0.063) and two significant differences were located in the other two texture (F2=4.888, P2=0.035; F3=42.994, P3=0.000), respectively. An ANOVA test comparing the influence of the four backgrounds revealed no significant difference for any of the texture options (F1=0.078, P1=0.972), (F2=0.3338, P2=0.798), and (F3=0.000, P3=1.000). The influence of background colour on the three texture parameters is illustrated in Figure 3.7.

3.3.1.4 Orientation of the label.

Considering the orientation of labels, two levels were considered. First, the orientation of the overall label was considered. Three sets of orientation were presented simultaneously to the participants: horizontal; always under an angle in correspondence to the shape of the object (tilted); horizontal if possible, otherwise under an angle which corresponds to the general orientation of the polygon (mixed). The preferences linked with orientation options were compared with the four background types. Second, the orientation of the label syllables was compared as straight versus italic.

The recorded preference values (in %) for the orientation of the overall label are listed in Table 3.2. The same trend was recorded for the four backgrounds. The majority of the participants tended to prefer the horizontal orientation (>60%); the two other orientation options were almost equally preferred. Most slightly higher preference was measured for the 'mixed' option, except for the blank background. User preferences for the three tested orientations were significantly different (one-way ANOVA, F=324.476, P=0.000). The similarity of responses towards the different background colours was also statistically confirmed (one-way ANOVA, F=0.39, P=0.996).

Table 3.2: Users' preference on label orientation related to background colour

Map design	Horizontal	Mixed	Tilted
Blank Map	66	16	18
Grey Scale	64	22	14
Hot Colour	62	22	16
Cold Colour	66	24	10

The obtained preference data concerning the label syllables is illustrated in Figure 3.8, in which the trend of the user preferences regarding the labels' orientation was compared with font size (8, 10, 12, and 14 dots). Generally, the straight version of the labels was preferred over italic labels. Using one-way ANOVA, a significant difference was found (F=29.985, P=0.000) between straight and italic labels along the four backgrounds. No significant difference was found in user preferences along the four background colours (F=0.418, P=0.741) when testing the second level of orientation. Users' preference for the straight versus the italic orientation variable of the four tested sizes regarding the influence of background colour is illustrated in Figure 3.9.

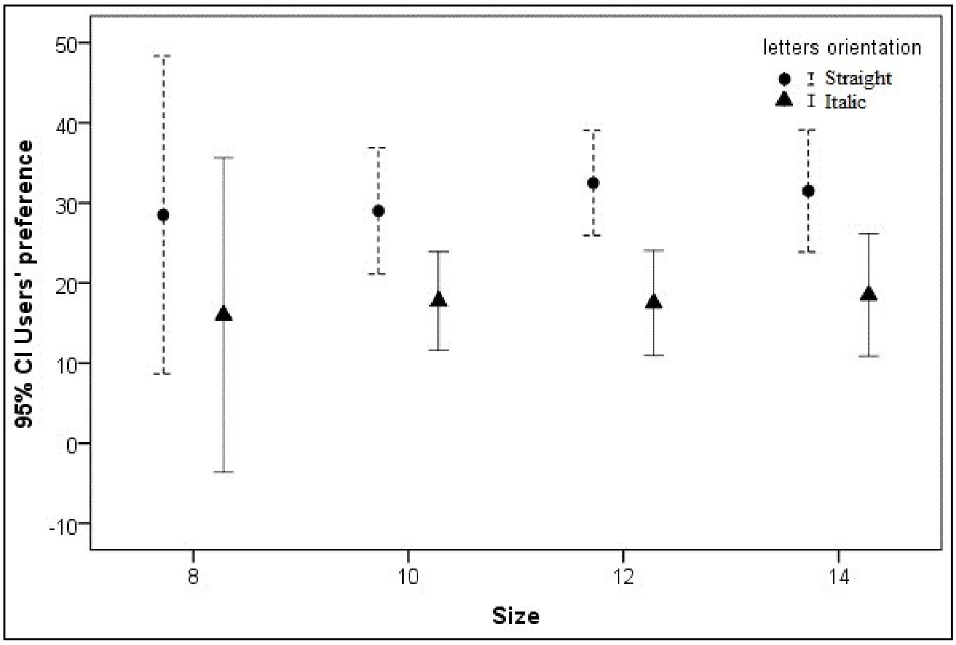


Figure 3.8: The mean user's preference on label orientation (straight versus italic) for the four backgrounds' colour tests

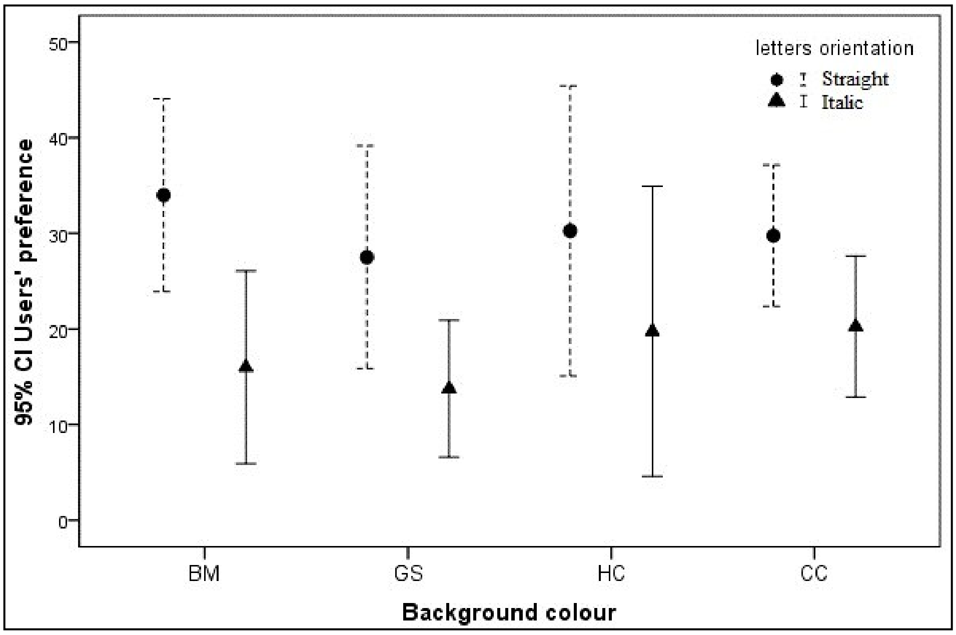


Figure 3.9: The influence of background colour onto the mean preference of label orientation (straight versus italic)

3.3.2 Labels associated with point data

The points were distributed randomly over the map image to mimic a realistic scene. Therefore, the associated name labels were distributed accordingly. A combination of several visual variables was applied to these labels to create different categories and hierarchical levels. Labels are flexible symbols because of a broad array of visualisation options in which they can be applied to the same placement. This finding can be divided into two levels: the same layout for all labels on the map face, and variations in the layout of individual labels on the same map face.

3.3.2.1 Size and shape variations

The stimuli used the case style with point data as they serve many functions of the cartographic text. The simplest combination of size and shape is the link between the variation in case style and font family. Three sets of the case styles were tested for Arial and Times New Roman: all characters in lower case, the first character in upper case, and all characters in upper case. User preferences are listed in Table 3.3. Therefore, a comparison was made of the preference data regarding the two font types across the four backgrounds (Arial versus Times New Roman, TNR): one-way ANOVA showed a significant difference in users' preferences in the case style variations across Arial and Times New Roman (F=535.297, P=0.000), Figure 3.10 illustrates this difference. Using one-way ANOVA, the preferences for these case styles were compared along the four background types. The obtained data showed no significant influence on the background colour on users' preference for a certain case style (F=0.000, P=1.000).

Table 3.3: Percentage of users' preference according to case variation and font type

Case style	**All letters are lower case**		**First letter is upper case**		**All letters are upper case**	
Font type	**Arial**	**TNR**	**Arial**	**TNR**	**Arial**	**TNR**
Blank Map	88	12	68	32	80	20
Grey scale	96	4	82	18	80	20
Hot Colour	86	14	90	10	78	22
Cold	82	18	84	16	80	20

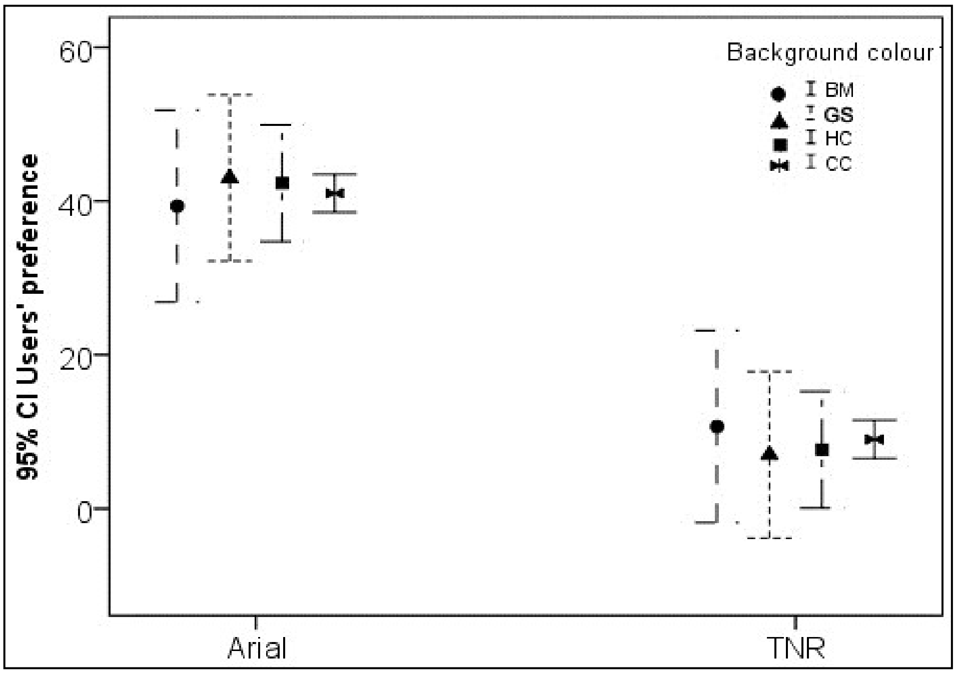


Figure 3.10: The influence of background colours onto the mean user's preference of labels' shape including case style variations

3.3.2.2 Texture of labels

The hierarchy of labels (and their associated objects) can be expressed through a combination of multiple visual variables. A first option is to vary the size of the labels. Because a bold representation of labels increases their size, differences in preference may occur. Therefore, a collection of two sets of maps were presented to the user: one without bold labels and one with all labels presented in bold. The labels on the two maps were both visualised in Arial or Times New Roman, resulting in four maps that were tested pair-wise. In a first sub-collection, only two levels of hierarchy were presented in the map image, using label sizes of 10 and 12 dots. A comparison of users' preference along the four backgrounds resulted in no significant differences (one-way ANOVA, F=0.07, P=0.97). The second sub-collection has four maps with labels varying between three sizes (8, 10, and 12 dots) to express hierarchy. The maps, combining bold and normal representation with Arial and TNR, also showed no significant difference along the four backgrounds as measured by one-way ANOVA (F=0.42, P=0.73).

Next, user preferences towards these textures were examined, with the maps used previously mixed to create new pairs. In this case, the maps with two levels of hierarchy (10 and 12-dot label sizes) were compared with maps that have three levels of hierarchy (8-, 10-, and 12-dots label sizes). Here, a distinction was made between the Arial and Times New Roman fonts. Using a one-way ANOVA, the results over the four background types were analysed, but no significant difference was found in the users' preferences between the Arial font (F=0.22, P=0.88) and the Times New Roman font (F=0.04, P=0.98).

3.3.2.3 Combining the variations of labels

Finally, a combination of multiple visual variables was applied to the labels within one map: different fonts (shapes), case style variations, boldness, and three size variations (8, 10, and 12 dots). In each map, three levels of hierarchy were presented using these variables. Accordingly, four different maps were formed and examined in pairs, resulting in six comparisons. No significant difference in user preferences regarding these four maps was found of the background colour (one-way ANOVA, F=0.62, P=0.60).

3.4 Discussion

This study shows that the combination of the typographical design, which involves the application of visual variables to introduce different labelling functions with the maps' background colour introduces no influence on users' preference regarding these visual variables. The findings show a consistent agreement of users' preference regarding the four map backgrounds (blank background, grey scale background, hot-coloured background, and cold-coloured background). As a consequence, the typographic design of screen maps should not be adapted according to the map's background colour. However, it should be noted that the perception of the horizontal relationship between typographic objects and their design, discussed by Steiniger and Weiber (2007), is not influenced by background colours. In addition to that, this result contradicts the findings of Bartz (1970), who found that the figure-ground and the surrounding relationships show more of an influence than the typographic design. However, her results described users' ability to locate varied labels.

Previous studies concerning the typographic design described its influence on readability. Unlike Arditi and Cho's (2005) conclusion from their experiment on the efficiency of normal reading, where the serif font was more efficient than the sans serif font, users' preferences of reading map labels were oriented towards the use of sans serif fonts for the three case styles (all letters in lower case, first letter in upper case, all letters in upper case). In addition to that, the users' preferences for the sans serif fonts Arial Narrow Bold, Arial Bold, and Comic Sans SM are significantly different from Feldmann and Kreiter's (2006) findings in terms of efficiency measurements.

It is noteworthy to stress that users' preference for larger sizes decreases when using bold labels. In addition to that, it should be mentioned that using different textures of labelling -with variation in italic, bold, and narrow styles- carries significant differences in the preferences of uses, who appear to prefer a combination of less visual variables.

Considering the orientation of labels, users significantly preferred the horizontal orientation for all background colours; this finding is likely due to the high-frequency use of horizontal orientation, especially on small-scale maps and the habitual orientation of normal reading as well. This result is consistent with Deeb *et al.* (2012) finding when they investigated the preference of orientation for two users' groups considering expertise and gender variations. Furthermore, the straight orientation was highly preferred over the italic orientation.

Regardless of whether the design of labels over maps considers two or three levels of hierarchy and whether this hierarchy is visualised using an Arial font or Times New Roman font, no significant difference could be attributed to the relation between the hierarchy composed of two or three sizes (normal or bold).

3.5 Conclusions and Future work

Maps are composed of several different elements (points, lines, areas, and labels) that are all symbolised using (one or a combination of) visual variables. The application of these variables has an influence on how the object is perceived but also affects the perception of neighbouring objects. In particular, the use of colours has an influence on adjacent objects. The perception of the cartographic elements thus depends on the overall design of the map, which in turn is dependent on its type, function, and content. Labels can be considered the fourth object type in map design and are commonly placed on top of all other layers of information. Consequently, the visualisation of this background information influences the perception of the labels.

The goal of this study was to detect whether different background types have a profound impact on map users' preferences regarding the visualisation of the labels themselves (their typography). For completeness, two types of objects were considered: areal objects and point objects. Different labelling options were created and compared, allowing users to express their preference given digital maps. These labelling options include: size, font, case styles, bold, italic, and combinations therein. These maps were integrated in four sets of similar questionnaires. The difference of these questionnaires was the background colour used: blank background, grey background, hot-coloured background, and cold-coloured background. This design enabled a comparison of the results and the detection of possible influences of these background colours. The results obtained during this study can be summarised as follows:

- for each of the variations in typography, none of the tested variations showed a significant difference in user preferences related to the background colour of the maps. However, the experiment conditions showed that the grey scale background and hot-colour backgrounds had a significant difference on user preferences between normal and bold labels;
- the preference of different shape characteristics considering three escalating categories of formality, serif, sans-serif, serif and spacing, and finally sans serif and spacing, showed a tendency towards the moderate visual variation of each category. This was not the case for formality: users preferred fonts with a high level of formality;

- the horizontal orientation was the most preferable direction to position labels. A horizontal orientation enables the constant reading of labels despite potentially exceeding the available space. However, it has the disadvantage that the overall orientation of the associated (areal) object is not stressed, potentially leading other objects be overlapped. The italic presentation of labels was significantly less preferred than a straight presentation;
- the most preferred font type was Arial in a pair-wise comparison with Times New Roman. This result was valid for the three case designs (all lower case, first letter in upper case, all upper case);
- comparing the application of multiple typographic options resulted in no significant differences in the preference values. Therefore, further investigation concerning individual texture preference was required.

It can be concluded that the background colour had no significant impact on users' preferences regarding several typographic options. This finding indicates that the typography of maps should not be adapted when changing the background colour in digital maps. In thematic maps, the same typography can thus be used with a wide range of different colour ramps applied to choropleth maps. Moreover, topographic maps can present a wide range of different region types, which must be visualised accordingly. These different visualisations will consequently not influence users' preference towards the typography of the labels. This study does not fully cover the visual variables introduced by Bertin (1967). More research must be conducted to test the application of colours on the cartographic labels themselves.

A next step in the research will be the assessment of the efficiency of the user when different typographic options are used on the map. These efficiency measurements may or may not correspond to the preference values of the users; something the user finds aesthetically attractive is not de facto 'better' for the user. Efficiency could be measured using response time measurements when users must complete a task on the map. Furthermore, the level of users' experience may be crucial when testing efficiency.

3.6 References

Arditi, A. & Cho, J. 2005. Serif and font legibility. Vision Research, 45 (23), pp.2926-2933.

Aykin, N. M. 1989. Individual Differences in Human-Computer Interaction. Computers & Industrial Engineering, 14(1-4), pp.614-619.

Bertin, J. 1967. Semiology of Graphics: Diagram, Networks, Maps, Translated by William J., Berg,2011, Esri press, California.

Bartz, B.S. 1970. An Analysis of the Typographic Legibility Literature, Assessment of Its Applicability to Cartography, The Cartographic Journal, 7(1), pp.6-17.

Brewer, C. A. 1996. Guidelines for Selecting Colors for Diverging Schemes on Maps. The Cartographic Journal, 33(2), pp.79-86.

Cleveland, W. S. & McGill, R. 1983. A Color-Caused Optical Illusion on a Statistical Graph. The American Statistician, 37(2), pp.101-105.

Deeb, R., Ooms, K., & De Maeyer, P. 2012. Typography in the Eyes of Bertin, Gender and Expertise Variation. The Cartographic Journal. 49(2), pp.178-185.

Fairbairn, D. J. 1993. On the Nature of Cartographic Text. The Cartographic Journal, 30(2), pp.104-111.

Feldmann, H-U. & Kreiter, N. 2006. Neuaufbau der schweizerischen, Landeskarte: Inhalt und Kartengrafik. Kartographische Nachrichten, 3, pp.115–121.

Garlandini, S. & Fabrikant, S. I. 2009. ‘Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization’, In: Hornsby KS *et al.* (eds) Lecture Notes in Computer Science, Spatial Information Theory, 5756. Springer-Verslag, Berlin Heidelberg , pp.195-211.

Granger, G.W. 1955. An Experimental Study of Colour Preferences. The Journal of General Psychology, 52(1), pp.3-20.

Imhof, E. 1975. Positioning Names on Map. The American Cartographer, 2(2), pp.128-144.

Karssen, A.J. 1980. The Artistic Elements in Map Design. The Cartographic Journal, 10(2), pp.124-127.

Kraak, M. J. & Ormeling, F. 2010. Cartography, Visualization of Spatial Data, 3rd edition. Pearson Prentic Hall, Harlow.

MacEachren, A. M. 1995. How Maps Work. Representation, Visualization, and Design, Guilford Press, NY.

Monmonier, M. 1996. How to Lie with Maps, The University of Chicago Press, Chicago.

Nielsen, J. 1993. Usability Engineering. Morgan Kaufmann, San Francisco.

Ou, L., Lou, M. R., Woodcock, A. & Wright, A. A. 2004. Study of Colour Emotion and Colour Preference. Part ш: Colour Preference Modeling. COLOR research and application ,29(5), pp.381-389.

Peterson, M. P. 2003. Maps and the Internet. Elsevier, Oxford.

Rubin, J. & Chisnell, D. 2008. Handbook of Usability Testing. How to Plan, Design and Conduct Effective Tests, 2nd edition, Wiley Publishing, Indianapolis.

Steiniger, S. & Weibel, R. 2007. Relations among Map Objects in Cartographic Generalization. Cartography and Geographic Information Science, 34(3), pp.175-197.

Wood, M. 1993. The Map-Users' Response to Map Design. The Cartographic Journal, 30(2), pp.149-153.

Yu, D., Akau, M.M.U & Chung, S.T.L. 2011. The mechanism of word crowding. Vision Research, 52(1), pp.61-69.

4

Evaluating the Efficiency of Typographic Design: Gender and Expertise Variation

Modified from: Deeb, R., Ooms, K., Vanopbroeke, V. & De Maeyer, P., 2013. Evaluating the Efficiency of Typographic Design: Gender and Expertise Variation. The Cartographic Journal. 51(1), pp. 75-86.

Preface

Although the efficiency of label placement algorithms has been studied extensively, few studies considered the influence of the label designs on the efficiency of map readers. Labels are one of the most important elements on the map as they can provide more information than other symbols can. The design of the labels does have to stress the theme, shape, and functionality of the associated objects, which results in a more efficient interpretation of the map content by the user. How the label designs can enhance the map readers' efficiency (and thus the quality of the maps themselves) is the main objective of this study. A user study was conducted in which the participants were asked to locate a target label on a map. Different label designs were implemented across the trials. The participants' reactions times were registered to measure their efficiency and statistically analysed using a one-way ANOVA. Two different users' characteristics were considered: gender and expertise. Related to the size, shape, orientation, and texture of the labels, a number of significant differences ($P<0.05$) and trends were located. Differences in efficiency between males and females, on the one hand, and between novices and experts, on the other hand, were also described statistically. Consequently, recommendations can be formulated regarding the design of labels in order to obtain more efficient maps, keeping in mind the map users' characteristics.

4.1 Introduction

Cartography comprises a systematic symbolization of the features in the environment, nevertheless the addition of labels is crucial to finalize the cartographic product. Gerber (1981) stressed the importance of labels by identifying three subsequent levels for a successful cartographic interpretation ranging from the simplest to the most complex: the perception-recipe level (pictorial), the label level (pictorial-verbal), and the 'other knowledge about' level (verbal). In the midterm level, the label level, the map reader gains knowledge about the name of the object he tries to interpret. Consequently, labeling objects on maps contributes significantly to the users' interpretation process. This facilitates the link between the pictorial and verbal level, which is critical to interpret the map correctly. Imhof (1975) demonstrated that map lettering is a linguistic, practical, technical, and esthetic issue. Regarding the use of these labels on maps, two main issues can be identified. The first issue concerns the placement of the labels, both on the algorithmic level known as automated placement, and on the cartographic level (quality of the placement) described by Imhof (1975) as good or poor placement. The second issue concerns the actual design of the labels: how should they be presented in order to create maps with a higher quality and which can thus be interpreted more efficiently by the map users.

4.1.1 Label Design

The design, or representation, of map labels is strongly connected to the function and type of associated feature divided into position, linear, and areal designation (Imhof, 1975). Kraak and Ormeling (2010) differentiated between a primary label function of providing the geospatial addresses and a secondary function of indicating the nature of the represented object. For this they set rules to indicate hierarchal and nominal differences by changing the variables in which the label is presented. They proposed using the variation of boldness, size, spacing, colour value, and/or case style to represent hierarchal differences. They also proposed the use of color, shape, and straight script versus italic script to represent nominal differences. In addition to that, Krygier and Wood (2011) also set some functional rules for using type as a map symbol such as using different typeface to indicated different qualitative data and using size and type weight to present order. Deeb *et al.* (2012) linked the use of visual variable to the perceptual characteristic of label design (associativity, selectivity, order, and quantity). The density of names on a map depends on the content and purpose of the map. In order to facilitate the interpretation of the visual information on maps, Imhof (1975) set rules to avoid type accumulation, type overlapping, wrong orientation, and type spreading, then he demonstrated the drawbacks of bad designs. Peterson (2009) discussed the use of uppercase and lowercase and suggested that lowercase letters are read easily because of the different height of various letters resulting in a specific shape for each letter; meanwhile the uppercase letters have the same height and global shape which make them harder to read. Bartz (1970a) pointed out that applying typographic findings reported in non-cartographic literature is acceptable because of the similar functional task that cartographic labels deliver. In order to set empirical rules, Bartz (1970b) investigated typographic variables of both shape and size on paper maps to evaluate their influence on search time for both individual label and the complete searching task. Different design criteria were also investigated by Phillips (1981) as he tested the influence of character design regarding the word's initial letter, the word length, and the word shape. By using eye tracking method, he concluded that these elements affect the fixation time for the target names. The association with the typographic design function and the readability of the text on the map was discussed by Fairbairn (1993).

Wood (2000) described the functions of label design and set rules for principle positioning of lettering according to their own function, illustrating the optimal typographic design according to the shape of depicted features. To investigate the influence of the labels' shape on screen map readability, Feldmann and Kreiter (2006) conducted a controlled test on three sans serif fonts (Arial, Univers, and Frutiger), but they found no significant difference between the readability of the three of them. Before that, Arditi and Cho (2005) investigated the legibility of serif and sans serif fonts on a plain text and concluded that serif fonts is slightly more legible than sans serif fonts. The readability of labels is expected to be affected by the typographic design (van den Worm, 2001: 87-107). He proposed that bold typography could improve readability but this may clash with desirable perception characteristics. Ever since, no empirical proofs supported or disapproved this proposal.

4.1.2 Map Usability Testing

The development of the cartographic products was combined with a growing interest of assessing their usability; including map effectiveness, map efficiency, and map satisfaction (Faulkner, 2000). The usability of interactive mapping tools was tested by Andrienko *et al.* (2002) where they tested the tool learnability, memorability, and user satisfaction. Their study concluded that users were able to adopt the new ideas for map interactivity and manipulability. Nivala *et al.* (2008) tested the functionality and the features of web mapping sites, they found out some usability problems related to search operation, user interface, map visualization, and map tools. To avoid such problems they suggested some design guidelines for web mapping design. Effectiveness of map design methods were tested by MacEachren (1982) when he examined the effectiveness of both choropleth maps and isopleth maps for direct data acquisition and pattern memory, exploring the role of complexity in their design. He concluded that the isopleth maps are more effective in the term of memorizing general pattern. The design variable were also tested by Garlandini and Fabrikant (2009), they reached empirical results suggesting that size is the most efficient and effective visual variable to detect changes under flicker conditions and orientation proved to be the least efficient and effective visual variable among size, hue, value, and orientation. Users' preference (satisfaction) of label design associated with point and areal data was tested by Deeb *et al.* (2012). They applied Bertin's (1967) visual variables on labels and investigated the users' preference towards different label designs and concluded some aesthetics design criteria. In addition to that, they indicated significant differences between user groups which were tested: novices versus experts and females versus males.

Bearing in mind that the aesthetics of the label design, tested through users' preference, does not always reflect an optimal design. The main goal of a cartographic product is communication: to get a message across as efficient as possible. With efficiency, the facility with which a user can interpret the content is meant. Therefore, it is crucial to complete the aesthetics measurements with information about the efficiency of certain label designs towards different map users. In order to measure the efficiency with which the map content is interpreted, reaction time measurements are often used. The latter issue is the main concern of the user study described in this chapter.

Quantifying users' performance can be done either by the numbers of achieved tasks or the time user takes to complete specific task (Nielsen, 1993). Measuring reaction times (RT) is often used to answer questions regarding the users' cognitive process. A fundamental method to measure performance is measuring objective time to complete a task (Rubin and Chisnell, 2008). Measuring reaction times does not only reflect the mental processing speed, but is also linked to how fast the information is retrieved. The users reaction times consist out of the total time needed to execute an assignment including the cognitive processes linked to it: analysing, storing, recoding, and subsequently using the raw data. In this chapter the users' reaction time to find a certain label on a map, linked with a number of label designs, is used as a measurement of their efficiency. The time consumed to find the target labels reflects the effect that different label designs have on the users' efficiency to read the map contents. Typically users pass by four stages to locate their targets. First stage is identifying the target, then memorizing the name, third is to adjust the typography of the target to the typography used in the map and finally searching to find the target. The structure of the study ensures that all users can pass through these stages without any interfering elements that could cause a bias to the results. Due to the fact that user characteristics should be captured and taken into account during the design (Haklay and Nivala, 2010), the study has both a between- and within-user design as the efficiency of different label designs were tested, considering different user characteristics (gender and expertise). The design of the user study is described in detail in the next sections.

4.2 Study Design

4.2.1 Participants

Two types of user characteristics are studied in the experiment, gender and expertise, which is reflected in the selected participants. The novice group included 25 participants with an average age of 16.4 years. This group consists of pupils at the secondary school level and, consequently, with no previous training in cartography. The expert group consisted of 25 participants who work with maps on a daily basis and have at least a master's degree in geography. The average age within the expert group was 29.5 years. Of the 50 participants, 25 were female and 25 male. The average age of both females and males was 23.

4.2.2 Task and Stimuli

A series of forty maps was presented to each participant in an online questionnaire. Both areal and point data were integrated in the map design, represented by thematic and topographic maps. Two examples of such maps used in the experiment are illustrated in Figure 4.1. For each map the participants had to locate a target label in the map image. They were instructed to click on the label in the map image when it was found, which resulted in a reaction time measurement (time interval, in milliseconds, between the display of the map and the mouse click). After the mouse click, a new map and new target label was displayed. These instructions were formulated on the first screen of the test. After the completion of the forty maps, the participant had to fill in a post study questionnaire. Gender, age, level of expertise, and other personal characteristics were registered in this questionnaire.

The target label was displayed above each map and was depicted in a neutral font (OCR A Extended) that was not used in the map image itself. The length of the target labels was kept constant: 6 or 7 letters. The names were carefully chosen regarding their general shape, taking into account, among others, their cap height, x-height, and ascenders. The location of the target label within the maps was chosen carefully, based on eye movement analysis from Ooms *et al.* (2012). Visual search strategies can also vary between users as Lleras and Mühlenen (2003) clarified. They discovered that different search approaches were followed. The search strategies were divided into either a systematic approach, as users geometrically scanned the map (top to bottom, bottom to top, right to left, and left to right), or an intuitive approach as the users randomly searched the map image. In order to level out these search approaches, target labels were distributed equally on all direction and covered the four corners (see Figure 4.2). Labels design on the map varied in size, shape, orientation, and texture, which were applied individually or combined. An extensive explanation of these visual variables can be found in the work of Bertin (1967) and their application on map labels was described by Deeb *et al.* (2012). Fictive labels were placed on the map to assure that participants would not recognize the location. This could bias the results due to previous knowledge: a user can retrieve the position of a label much faster if he is familiar with the region.

Figure 4.1: An example of two test screens, (a) point data labelling, (b) areal data labelling

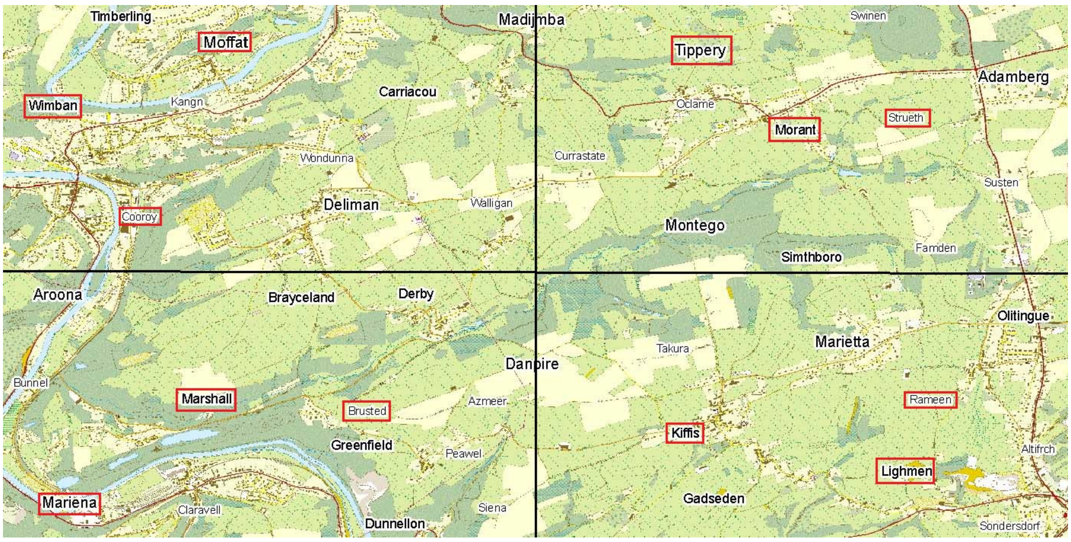


Figure 4.2: An example of label target distribution on the topographic map

To avoid biases in the measurement due to resolution and size differences, all participants completed their test on a flat screen with a 1280x1024 resolution. The latency of mouse registration time and the software registration time were taken into account. Furthermore, the test was run on hardware with similar properties and presented in the same browser (Google Chrome). Each map had a similar size (about 350 KB) to ensure equal loading times of it. Each participant followed the same order of maps in a sequence that lasted 10- 15 minutes. The same label was never asked more than twice and the studied variables were presented as such that maps related to the same variable were not presented directly after each other.

4.3 Results and Discussion

4.3.1 Data

All data were collected through an online questionnaire and stored in a database. However, not all participants completed the questionnaire until the end, and consequently, no personal characteristics could be gathered. In order to assure the consistency of the dataset, all incomplete records were removed from the database. Furthermore, out of 2000 measurements 54 outliers were detected in the reaction time measurements. All these outliers were larger than the mean value plus two times the standard deviation (M+2 SD; indicating the limit of the 95% confidence interval).These outliers could be explained by distractions of the participant during the experiment. If the participant is distracted at a certain point, the reaction times would show a steep increase. However, these values do not contribute to the goal of our .decided to remove these outliers from the dataset. No outliers less than [M-2 SD] were found, which support the theory of participants' distractions related to the other outliers.

4.3.2 Size of Labels

Using Arial font, four label sizes were integrated in the test: 8, 10, 12, and 14 points. In addition to that, a bold typography was applied to the previous label sizes, which even enlarges the characters. The users' reaction time measurements were collected and a one-way ANOVA test was applied to both groups of gender and expertise. In order to get an overview of the mean values and distributions of the registered reaction time, Figure 4.3 illustrates the box plots of each category and group. From this figure it can be derived that the reaction times do vary between the user groups, however no systematic trend can be observed for the four different sizes. The shape of the box plots represents the range of reaction time observations and it can be noticed that the range of the reaction time measurements related to a normal font was larger than these related to a bold font. This remains true for both gender and expertise sub-groups.

Table 4.1 lists the mean reaction times (M) and associated standard deviations (SD) for males and females, related to the different label sizes. This table shows a trend in the reaction time measurements related to the bold and normal design of labels: both user groups were numerically faster in finding the labels with a bold design in comparison to the normal label design. A one-way ANOVA test on reaction time measurements of normal size versus bold size for gender (females versus males) demonstrates this difference is highly significant (F=13.600, P=0.002). Table 4.1 also indicates decreasing reaction times when larger labels sizes were used for the males. However, the labels with size 12 points in bold caused a deviation in this trend. In the females' user group, more deviations from this trend were observed. A one-way ANOVA shows that the males' and females' reaction time measurements were very similar: only the reaction times related to locating labels with a size of 14 points (normal) were significantly shorter for the males (M=6354 ms) than females

(M=10599 ms). As only one value near to significant difference demonstrate the variation between the males and females for the size 8 bold. The most efficient label size seemed to correspond to bold 14 points for the males (M=4873 ms) and bold 10 points for the females (M=4794 ms).

Table 4.1: Mean reaction times (in ms) for females and males, related to label sizes

Size		Females		Males		ANOVA	
		SD	***M***	***SD***	***M***	***F***	***P***
8	**normal**	11564	7901	11464	8133	0.002	0.965
	bold	9670	4984	6976	4892	3.719	0.060
10	**normal**	11048	10294	12320	9931	0.260	0.613
	bold	4794	3320	5306	4033	0.221	0.640
12	**normal**	8094	5976	7099	5408	0.381	0.540
	bold	5626	3679	6201	3722	0.303	0.585
14	**normal**	10599	8533	6354	5036	4.161	**0.047**
	bold	6119	2808	4873	2460	2.732	0.105

The results, listed in Table 4.2, present an overview of the mean (M) reaction time measurements and standard deviations (SD) for the novices and experts. The last two columns show the results from the statistical comparison between the two groups of users, using a one-way ANOVA. This table shows a similar trend as with the males and females: all users (novices versus experts) could find bold labels faster than normal labels. Furthermore, this difference can be considered highly significant (F=10.670, P=0.006). This means that the bold design of labels was more efficient than the normal design towards these different user groups. Furthermore, the results from the ANOVA test revealed a similar level of efficiency between the two groups of users. However, two significant differences were measured between both user groups. Novices could locate the labels with a size of 8 points (normal) much faster (and thus efficiently) than the experts. In contrary, the experts could locate the labels with a size of 12 points (normal) more significantly efficient than the novices. The novices considered the bold labels with a size of 14 points as the most efficient design (M=5532 ms), however only one test showed a near to significant different value with a size of 10 points was noted (M=5594 ms). The experts could locate the labels with the latter design (bold, 10 points) most efficiently (M=4516 ms).

Table 4.2: Mean reaction times (in ms) for novices and experts, related to label sizes

Size		Novices		Experts		ANOVA	
		M	***SD***	***M***	***SD***	***F***	***P***
8	**normal**	9155	6031	13874	8978	4.758	**0.034**
	bold	7156	4684	9490	5271	2.737	0.105
10	**normal**	9293	5883	13954	10294	0.613	0.059
	bold	5594	4015	4516	3305	1.004	0.322
12	**normal**	9201	5697	5993	5260	4.279	**0.044**
	bold	6581	4020	5246	3236	1.672	0.202
14	**normal**	9164	7440	8084	7399	0.249	0.620
	bold	5532	2908	5437	2508	0.015	0.903

The experiment indicate that all users groups have a significant better performance when using bold size as their reaction time measurements were significantly less than those of normal size measurements. Matching reaction time measurements to the four tested font size did not show any systematic trend to predict the efficiency according the increase or decrease of font size. Nevertheless, it showed typical cases where both groups agreed with font size (both point size and boldness) and other cases of disagreement. Variations in size correspond to variations in performance according to both the gender group and the expertise group. The four tested size of Arial font (8, 10, 12, and 14) showed different trends but what is interesting about these trends is that all comparisons agreed on more efficient performance for bold design in the four tested sizes. In addition to that, the bold design

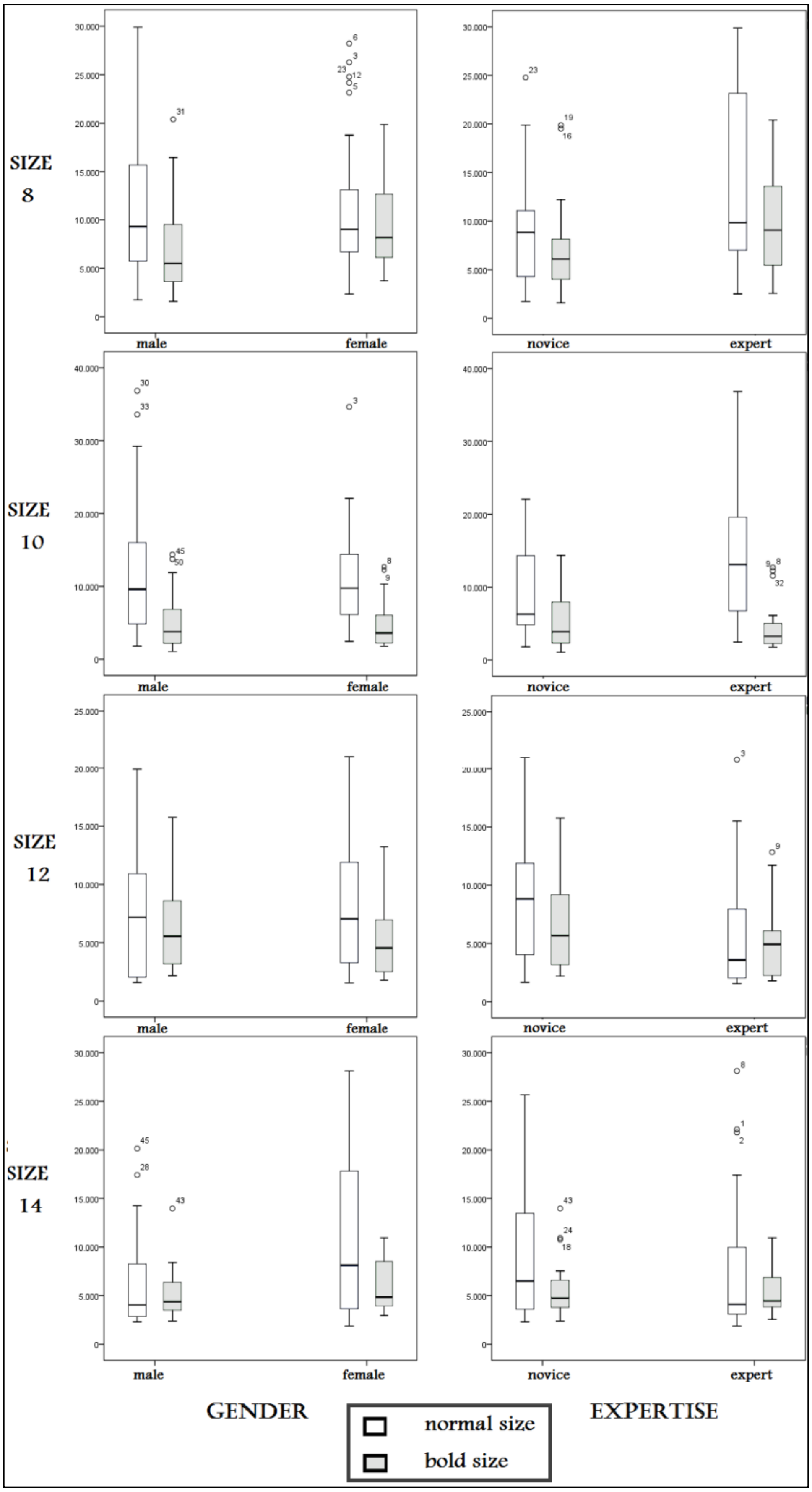


Figure 4.3: Box plots of the reaction time measurements (in ms) related to the labels' size, boldness, and user groups

did not witness any significant difference for the different point sizes and the different sub-groups. This results agrees with van den Worm (2001) who suggested that bold typographic design is more readable than the normal typographic design. In addition to that, it also contributes to what Deeb *et al.* (2012) presented in their research of users' preference towards label design as they found out that bold size design is more preferred than the normal size design.

4.3.3 Shape of Labels

The shape of label is mainly determined by its font: its typeface at a certain size. Theoretically, the world of shapes is infinite, which thus hold also true for fonts (Bertin, 1967). In this infinitive amount of fonts, different families can be identified, from which 'serif fonts' and 'sans serif fonts' are two of the most important ones. The efficiency of these two font families was tested using two of its most used (and known) typefaces, namely Arial (sans serif) and Times New Roman (TNR, serif). Both fonts were presented at a size of 10 points (normal) and on a blank background to avoid any disturbance of the surrounding elements and influence of colours on the user's cognitive load.

Table 4.3 lists the mean (M) reaction times of the males and females and 4.4 lists the mean reaction times of the novices and experts. The last two columns of both tables contain the results of the ANOVA tests between the sub-groups. These tests indicate that the efficiency with which the labels were located is similar in all user groups. Only one value near to significant difference was detected between novices and experts for Arial font.

Table 4.3: Mean reaction times (in ms) for females and males, related to the labels' shape

Font	**Females**		**Males**		**ANOVA**	
	M	***M***	***SD***	***SD***	***F***	***P***
TNR	6530	7021	5047	3899	0.140	0.710
Arial	11048	12320	9931	7408	0.260	0.613

Table 4.4: Mean reaction times (in ms) for novices and experts, related to the labels' shape

Font	**Novices**		**Experts**		**ANOVA**	
	M	***SD***	***M***	***SD***	***F***	***P***
TNR	7086	4292	6486	4758	0.210	0.649
Arial	13954	5883	9293	10294	3.744	0.059

Analysing the mean reaction times based using one-way ANOVA test showed no influence on the use of a serif or sans serif fonts on both gender (F=0.546, P=0.537) and expertise (F=4.238, P=0. 176). Both the between- and within-user reaction time study agreed on the efficiency of utilising serif and sans serif fonts. Further designs of serif and sans serif fonts on coloured backgrounds were undertaken to demonstrate the influence of case style variation on users' efficiency of both Times New Roman and Arial fonts: all lowercase, only first uppercase, all uppercase. The results of these six different designs are depicted in Table 4.5 (females versus males) and Table 4.6 (novices versus experts).

In both tables, the reaction time measurements seemed to be rather similar between the two user groups. Only one significant difference could be detected between males and females: females could locate the labels in the 'First letter uppercase, Arial' design significantly faster than the males (F=6.155, P=0.017). No general trend could be observed between the use of the Arial and Times New Roman font regarding the efficiency of the users. However, one remarkable trend was noticed in the 'all lower case' design. The largest reaction time measurements for all user groups were registered with the Times New Roman font, whereas the smallest reaction time measurements were linked with the Arial font. This would mean that this design is the least efficient when the Times New Roman font is used, and it is at the same time the most efficient when the Arial font is applied.

Table 4.5: Mean reaction times (in ms) for males and females, related to categories of labels using case styling

Categories	Font	Females		Males		ANOVA	
		M	***SD***	***M***	***SD***	***F***	***P***
all lowercase	**TNR**	11832	8094	13001	6469	0.138	0.575
	Arial	6233	2365	6587	2249	0.282	0.598
first uppercase	**TNR**	6738	3742	6875	3203	0.018	0.895
	Arial	6384	2683	9548	5783	6.155	0.017
all uppercase	**TNR**	7968	4120	8439	4400	0.141	0.709
	Arial	10092	7826	8482	4152	0.785	0.380

The expertise group showed two values near to significant when ANOVA test was applied to test the significant variation between the user sub-groups. Both values were recorded in the category of all letters lower case. For TNR font novices (M=10491ms) performed the task faster than experts (M=14342 ms) and this was near to significant difference (F=3.695, P=0.061).For Arial font experts (M=5844 ms) performed the task faster than novices (M=6976 ms) and this also was near to significant (F=3.060, P=0.087).

Table 4.6: Mean reaction times (in ms) for novices and experts, related to categories of labels using case styling

Categories	Font	Novices		Experts		ANOVA	
		M	***SD***	***M***	***SD***	***F***	***P***
all lowercase	**TNR**	10491	6448	14342	7666	3.695	0.061
	Arial	6976	2298	5844	2184	3.060	0.087
first uppercase	**TNR**	6857	3668	6752	3245	0.010	0.920
	Arial	7024	3709	8908	5500	2.016	0.162
all uppercase	**TNR**	7231	4055	9021	4265	2.101	0.154
	Arial	10043	7608	8528	4539	0.695	0.409

To obtain general conclusion about the efficiency of user groups towards the two fonts for the described design, a one-way ANOVA test was used to describe statistically the efficiency of TNR and Arial. The gender sub-groups showed no significant difference in their reaction time measurements of TNR and Arial in the three described case-style (F=0.952, P=0.352) and so did the expertise sub-groups (F=0.819, P=0.352).

Notwithstanding its general character, the results reported here would seem to indicate that the efficiency of the studied font family (serif, sans serif) did not affect the user reaction time measurements. However, the case styling of the labels in combination with a certain font family did show a significant influence on the user's efficiency between males and females. It is critical to establish rules for the shape design efficiency spatially when the design is used for different function. this topic was researched thoroughly preference wise by Deeb *et al.* (2012) when they tackled the case style, they concluded that 70-80 % of users preferred Arial font over Times New Roman. The participants' efficient performance on Arial contradicts Bartz (1970a), who suggested applying typographic findings that have been reported in the non-cartographic literature. Because this result disagree with Arditi and Cho (2005) who concluded from their experiment on plain text that serif fonts are more legible than sans serif fonts.The result considering the efficiency of case style could be explained as Peterson (2009) demonstrated; it is easier to read the lowercase letters because they have different height and thus different shapes, unlike the uppercase letters which all have the same height and makes the process of distinguishing letters longer.

4.3.4 Orientation of Labels

The orientation of the labels on a map can be considered at two levels: orientation of the whole label and orientation of the characters within the label. Studying the first level, labels were placed according to three approaches. First approach (Horizontal) considers placing all labels horizontally. The second approach (Tilted) implies placing the label under an angle, based on the shape of the object. In this case, label will stress the general shape of the object, as the labels are placed along its main diagonal. The third approach (Mixed) gives a higher priority to horizontally placed labels. Only when this is not possible, the labels will be tilted. These three types of label orientations were integrated in the test. The results of the mean reaction time measurements (M) and the statistical comparison between the different user groups are presented in and Table 4.7 (males versus females) and Table 4.8 (novices versus experts).

In the gender comparison, one significant difference was detected, related to the horizontal orientation of the labels (F=5.685, P=0.021). Females (M=6768 ms) could locate these names faster (and thus more efficiently) than the males (M=9785 ms). In the case of the novices versus experts comparison, also one significant difference was detected but this time for the tilted approach (F=6.314, P=0.015). The novice participants (M=8428 ms) seemed to be able to located the tilted labels more efficiently than the expert participants (M=14868 ms).

Table 4.7: Mean reaction times (in ms) for males and females, related to the labels' orientation

Orientation	**Females**		**Males**		**ANOVA**	
	M	***SD***	***M***	***SD***	***F***	***P***
Horizontal	6768	3213	9785	5188	5.685	0.021
Tilted	13460	10534	9626	7924	2.059	0.158
Mixed	13121	10546	13090	10080	0.432	0.514

Table 4.8: Mean reaction times (in ms) for novices and experts s, related to the labels' orientation

Orientation	**Novices**		**Experts**		**ANOVA**	
	M	***SD***	***M***	***SD***	***F***	***P***
Horizontal	7895	4208	8705	4916	0.367	0.548
Tilted	8428	5736	14868	11402	6.314	**0.015**
Mixed	12511	9899	13724	10686	0.936	0.338

No systematic trend could be observed in the data. Females and novices were numerically faster in locating the horizontally placed labels than with tilted and mixed orientation. However, the males and experts showed little difference in the mean reaction times between the horizontal and tilted orientation, and a larger difference with the mixed approach. However, none of the mean reaction times pointed to the mixed design as the most optimal label orientation (in terms of efficiency).

Within-users analysis and under the experiment condition, using one-way ANOVA did not show any significant difference of the reaction time measurements of the orientation design for both gender group (F=3.063, P=0.188) and expertise group (F=1.678, P=0.324). However, it would be interesting to extend this work to cover different functions of the orientation placement. This result agrees with Garlandini and Fabrikant (2009) as they found that the visual variable orientation proved to be the least efficient and effective visual variable. The three orientations did not cause any significant difference between the users sub-groups of expertise and gender. The trend of user performance in reaction time tasks agrees with the trend that Deeb *et al.* (2012) introduced, when they measured users preference towards the orientation of label design. Horizontal orientation was the most preferred orientation over the tiled and mixed orientation.

Table 4.9: Mean reaction times (in ms) for males and females, related to the use of italic

Italic	Females		Males		ANOVA	
	M	*SD*	*M*	*SD*	*F*	*P*
Size 8	7215	4230	7028	3109	0.029	0.866
Size 10	5141	2670	4683	3209	0.288	0.594
Size 12	4728	2454	5203	3387	0.319	0.575
Size 14	5581	1726	6261	2856	1.016	0.319

Table 4.10: Mean reaction times (in ms) for novices and experts, related to the use of italic

Italic	Novices		Experts		ANOVA	
	M	*SD*	*M*	*SD*	*F*	*P*
Size 8	6155	2944	7940	4109	2.765	0.103
Size 10	5649	3686	4234	1831	2.904	0.095
Size 12	5337	3460	4599	2321	0.775	0.383
Size 14	5930	2066	5883	2624	0.005	0.946

When considering the second level of label orientation, the individual characters are tilted. This is achieved by using an italic (Arial) font. The results of the normal font (straight), are listed in 4.1 and Table 4.2, and with those presented in italic are listed in 4.9 and Table 4.10. In order to obtain a structured overview, the box plots of these measurements are depicted in Figure 4.4. From this figure it can be derived that the mean reaction times are systematically higher when a label in a straight font had to be located. The mean reaction time measurements for the gender sub-groups indicated that labels in the italic design could be located significantly faster than in the straight design (ANOVA, F=21.549, P=0.000), and thus more efficiently. Similar results were found for the expertise sub-groups (ANOVA, F=15.841, P=0.001). This could be due to the attraction that italic causes as it is not the typical design of labels and usually used to make distinction of thematical purposes.

Using Arial font, four italic sizes were tested using ANOVA, which indicated no significant difference in both gender and expertise groups (see Table 4.9 and Table 4.10) as (P > 0.05). A trend was noticed; for the smallest size (8 and 10) males were faster than females meanwhile females were faster for the larger size (12 and 14). Novices were faster for the smallest size 8 and experts were more efficient than novices for the larger sizes (10, 12, and 14). Only one value near to significant was recorded and it was for size 10 italic in the expertise group (F=2.904, P=0.095). As italic showed significantly better users performance, it could be used as a powerful element to indicate importance of some map elements.

4.3.5 Texture of Labels

The design of the labels' texture has a wide array of choices as they are based on a combination of visual variables (Deeb *et al.*, 2012). Using textures, different categories of labels can be visualised by combining different label designs (one design for each category). These categories might correspond to different hierarchic levels in the labels (e.g. city versus village) or functional classes (rivers versus roads). Hierarchic levels can be presented by variations in the labels' size. In this study, three consecutive sizes were used (8, 10, and 12 points) to present three levels of hierarchy. Size of 12 points and 10 points were used to present two levels of hierarchy. Both levels of hierarchy were presented in a Times New Roman (TNR) and an Arial font. In Table 4.11, the mean (M) reaction time measurements for the males and females users are listed, together with their mutual statistical comparison using ANOVA. Table 4.12 displays these results for the novice and expert users and Figure 4.5 showes an example of 2 levels and 3 levels of hierarchy.

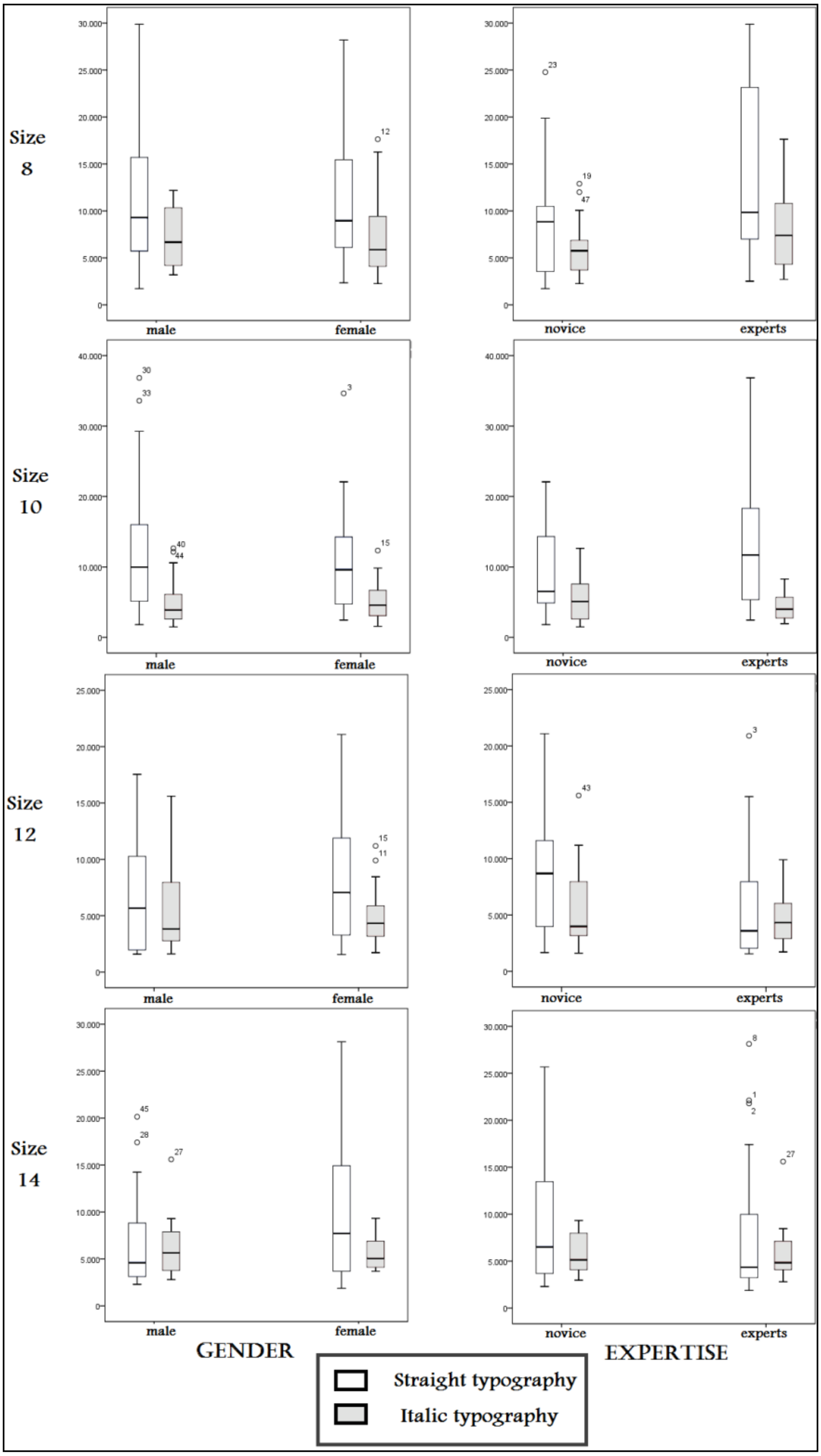


Figure 4.4: Box plots of the reaction time measurements (in ms) related to the labels' orientation, straight versus italic and user groups

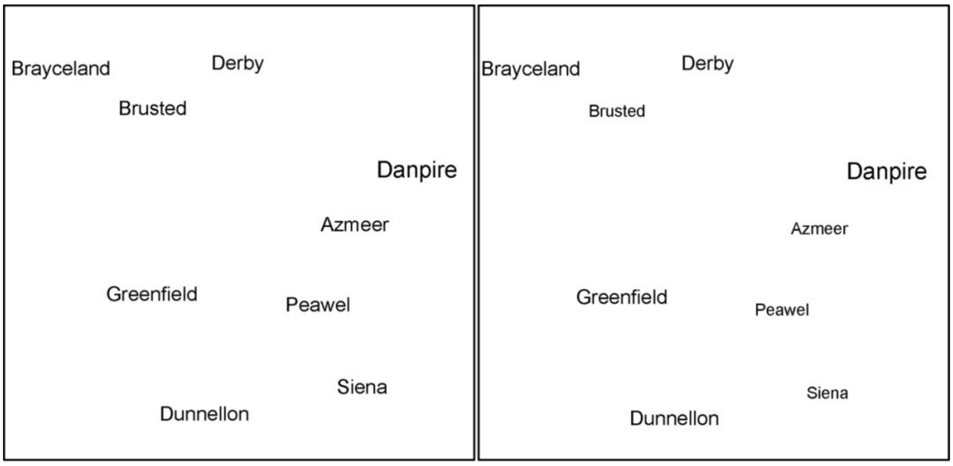


Figure 4.5: a comparision between 2 levels (left) and 3 levels (right) of hierarchy with Arial shape

The lowest reaction times were linked with three levels of hierarchy in Times New Roman. However, these measurements did not deviate much from these linked with the Arial font. Larger deviations were found between the designs with respectively two and three levels of hierarchy. Longer search times were registered when only a ‘2 Levels’ hierarchy was depicted. One exception on this was the mean reaction time measurements of the novices for locating labels when a ‘2 Levels’ hierarchy was depicted in an Arial font. This latter combination resulted in a mean search time which was rather similar to the one of the ‘3 Levels, TNR’ label design. This was not the case within all other user groups. One significant difference was located in the group of expertise for the ‘2 Levels, Arial’ (F=6.982, P=0.011) when novices performed significantly faster (M=7768 ms) than experts (M=13029 ms).

Table 4.11: Mean reaction times (in ms) for males and females, related to categories of labels using font sizes

Categories	**Font**	**Females**		**Males**		**ANOVA**	
		M	***SD***	***M***	***SD***	***F***	***P***
3 Levels	**TNR**	8040	4620	8797	4509	0.337	0.564
	Arial	9827	6031	9661	5113	0.011	0.918
2 Levels	**TNR**	11613	5651	12842	5589	0.573	0.453
	Arial	10282	6869	10629	8037	0.027	0.871

Table 4.12: Mean reaction times (in ms) for novices and experts, related to categories of labels using font sizes

Categories	**Font**	**Novices**		**Experts**		**ANOVA**	
		M	***SD***	***M***	***SD***	***F***	***P***
3 Levels	**TNR**	7818	4663	9010	4417	0.844	0.363
	Arial	9414	5764	10091	5404	0.180	0.674
2 Levels	**TNR**	11078	6070	13236	5022	1.813	0.185
	Arial	7768	4190	13029	8844	6.982	**0.011**

A one-way ANOVA test showed a statistical difference between both categories (2 Levels versus 3 Levels) when males and females were considered (F=10.196, P=0.019). However, there was not any statistical difference between both categories when expertise groups were considered (F=2.628, P=0.156). The gender group indicated that the labels in the 3 levels design could be located more significantly efficient than in 2 level design (see Table 4.11). Table 4.12 shows that there is no trend to the expertise sub-groups reaction time measurements which indicates the need for further research in this domain. No systemic trend was detected for the influence of complexity on users' efficiency. But for the gender sub-groups. There was a trend addressing the role of functional complexity; the higher the complexity (3 levels), the more efficiency the design is. This result disagrees with MacEachern (1982) conclusion on the influence of complexity effectiveness wise on data direct acquisition, considering isopleth and choropleth symbolization methods. Because he found out that no significant difference of users performance towards the two symbolization methods.

4.4 Conclusions and Perspectives

The efficiency of map labels involves much more than creating legible labels. In addition to the optimal label placement, the functional design of the labels should be provided (Fairbairn, 1993). Four types of label designs were included in a user test, which are used to indicate thematical, hierarchical, and shape of the cartographic information (Deeb *et al.*, 2012).

Comparing the reaction time measurements for both groups, bold size does not carry any significant difference for both gender and expertise group. However, a trend was noticed here: all users could locate the labels more efficiently when they were presented in bold as opposed to not-bold (normal). The efficiency of the users across the different groups seemed to be very similar, from which it can be concluded that the design of the labels should not be made for each user group specifically. However, generally speaking, all users seem to be able to locate names more efficiently when the size of the labels (in points) increases. This can be explained by the fact that the labels will become better readable with a larger size. This is compatible with the general rule that van den Worm sat (2001) saying 'text to be rasterised should not be smaller than 10 point'. Thus, it is not recommended to use the size 8 point for the mentioned reasons.

What is interesting about bold size is that it proved its efficiency over the normal size for both groups. Using bold size shows significantly faster reaction time than the normal size for gender group and so do for the expertise group. In addition to that, the findings recommend the use of bold as it was significantly more efficient than the normal size. This also corresponds to what van den Worm suggested (2001) of using bold font. Furthermore, it also matches the findings of Deeb *et al.* (2012) as they concluded that the use of bold font is the preferred use over the use of normal font. It should be noted that changing the font changes the type size (van den Worm, 2001), therefore these results are limited to the use of Arial font and could be extended to cover the sans serif font family based on Feldmann and Kreiter (2006) exploratory conclusions.

Variation in label shapes does not indicate any significant difference of using serif and sans serif font. In addition to that, there is no significant statistical difference between males versus females and novices versus experts except for the gender study of Arial first letter upper case. It is noteworthy that the results are more consistent when only one case-style is used. Although these results do not fully solve the arguments about the use of label shapes, it leads to a coherent description of serif and sans serif design when users' characteristics are paired in mind.

Orientation variable shows a significant difference located in the gender group for the horizontal orientation (females were faster). In addition to that, the reaction time indicates a significant difference in the expertise group for tilted orientation (novices were faster). Whereas no significant difference was detected for the mixed orientation in both gender and expertise study. The orientation of the label forming letters is expressed by italic. The use of the four sizes with italic do not show any significant difference in both gender and expertise groups. But it proved significantly its efficiency over the use of normal size and this could be explained by the attraction that the italic cause to the map users.

Hierarchical and thematical functions were used as a guide to design texture. Only one significant difference was deducted for hierarchy designs which use two level of hierarchy and Arial shape. When comparing the two designed hierarchy systems using ANOVA test, it shows that the hierarchy group composed of three levels is significantly more efficient than the hierarchy of two levels for the gender groups, meanwhile there is not any significant difference of both hierarchy when studying the performance of expertise group. When comparing these results with what Deeb *et al.* (2012) presented, it worth mentioning that users always prefer the use of two levels of hierarchy. However, this (exploratory) study may offer some insight into adequate hierarchy presentation in term of users' efficiency. Further analyses need to be undertaken to indicate the optimal hierarchal design for different labelling purposes.

These conclusions can be employed when designing web sites with cartographic product or web sites for cartographic use. If the web site is open to all visitors, designers shall avoid using the variables in which significant differences between users were located. They can use variables where results were more consistent such as bold and italic. For more efficient design, designer can customise a log in to the user interface to determine users' characteristics such as gender and level of experience. Consequently, they can use the most efficient variable for each function and group as described earlier. This can also be employed when designers know the audience characteristics; such as making maps and atlases for research use where they can use the results of expert reaction time which indicated the most efficient variable. Such products can also help students at schools by making maps for gender differences.

Future work of tracing the performance of map audience according to the label design variation is a crucial for both the map users and the efficiency of labels. Different age group response to label design will be studied as well as different map use which seems crucial to accomplish this work. In addition to that, labelling linear features and the properties of their design will be tackled. Furthermore, the efficiency of other alphabetical systems such as Arabic and Chinese will be studied and thus, a comparison with the Latin design finding could be made. Moreover, the visual variable of colour and value will be involved in label design with a suitable map use and suitable labels' function.

4.5 References

Andrienko, N., Andrienko, G., Voss, H., Bernardo, F. Hipolito, J., & Kretchmer, U. 2002. Testing the Usability of Interactive Maps in Common GIS', Cartography and Geographic Information Science, 29(4), pp.325-342.

Arditi, A. & Cho, J. 2005. Serif and font legibility. Vision Research, 45(23), pp.2926-2933.

Bertin, J. 1967. Semiology of Graphics: Diagram, Networks, Maps, Translated by William J. Berg, 2011.Esri press, California.

Bartz, B.S. 1970b. Experimental Use of The Search Task in The Analysis of Type Legibility in Cartography, The Cartographic Journal, 7(2), pp.103-113.

Bartz, B.S. 1970a. An Analysis of the Typographic Legibility Literature, Assessment of Its Applicability to Cartography, The Cartographic Journal, 7(1), pp.6-17.

Deeb, R., Ooms, K., & De Maeyer, P. 2012. Typography in the Eyes of Bertin, Gender and Expertise Variation, The Cartographic Journal, 49(2). pp.176-185.

Fairbairn, D.J. 1993. On The Nature of Cartographic Text , The Cartographic Journal, 30(2), pp.104-111.

Faulkner, X. 2000. Usability Engineering. Palgrave, NY.

Feldmann, H-U. & Kreiter, N. 2006. Neuaufbau der schweizerischen, Landeskarte: Inhalt und Kartengrafik, Kartographische Nachrichten, 3, pp.115–121.

Garlandini, S., & Fabrikant, S.I. 2009. Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization'. In: Lecture Notes in Computer Science, Spatial Information Theory, Hornsby, K.S. *et al.* (eds), 5756, pp.195-211, Springer-Verslag, Berlin Heidelberg.

Gerber, R.V. 1981. Competence and Performance in Cartographic Language, The Cartographic Journal. 18 (2), pp.104-111.

Haklay, M. M. & Nivala, A. M. 2010. 'User-centred Design', In: Interacting with Geospatial Technologies, Haklay, M. M.(ed), chapter 5, pp. 91-106, John Wiley& Sons. Chichester.

Imhof, E. 1975. Positioning Names on Maps, The American Cartographer, 2(2). pp.128-144.

Kraak, M. J. & Ormeling, F. 2010. Cartography. Visualization of Spatial Data, 3rd edition, Person Prentic Hall, Harlow.

Krygier, J. & Wood, D. 2011. Making Maps, A Visual Guide to Map Design for GIS, 2nd edition, The Cuilford press, NY.

Lleras, A. & Von Mühlenen, A. 2003. Spatial Context and Top-Down Strategies in Visual Search, Spatial Vision, 17(4). pp.465-482.

MacEachren, A. M. 1982. The Role of Complexity and Sympolization Method in Thematic Map Effectiveness, Annals of the association of American Geographers, 72(4), pp.459-513.

Nielsen, J. 1993. Usability Engineering, Academia Press, San Francisco.

Nivala, A. M., Brewster, S. & Sarjakoski, T. 2008. Usability Evaluation of Web Mapping Sites The Cartographic Journal. 45 (2), pp.129-138.

Ooms, K. Andrienko, G., Andrienko, N., De Maeyer, P. & Fack, V. 2012. Analysing the Spatial Dimension of Eye Movement Data Using a Visual Analytic Approach, Expert Systems with Applications, 39 (11), pp.1324–1332.

Peterson, G. N. 2009. GIS cartography : a guide to effective map design, Taylor & Francis, Florida.

Phillips, R.J. 1981.Search for Target in a Random Arrangement of Names: an Eye Fixation Analysis, Canadian Journal of Psychology, 35(4), pp.330-346.

Rubin, J. & Chisnell, D. 2008 Hand Book of Usability Testing: How to Plan, Design, and Conduct Effective Tests, 2nd edition, Wiley, Indianapolis.

van den Worm, J. 2001. 'Web Map Design in Practice', In: Web Cartography, Kraak, M.J. & Brown, A. (eds.) Chapter 7, pp.87-107, Taylor & Francis, London.

Wood, C.H. 2000. A Descriptive and Illustrated Guide for Type Placement on Small Scale Maps, The Cartographic Journal. 37 (1), pp.5-18.

5

Cartographic Text, the Preference of Label Design, and Cross-Cultural Comparisons

Modified from:
Deeb,R., Bonchev, S., Bandrova, T., & De Maeyer, P. 2014. How to Design Users' Friendly Multilingual Maps? Cyrillic and Latin Labels. 5th International Conference on Cartography and GIS. Riveiera, Bulgaria.
Deeb,R., Kurban,A., Ablikim, A., & De Maeyer, P.2015. From Monolingual to Multilingual Mapping: What is users' preference? 27th International Cartographic conference, Rio de Janeiro, Brazil (accepted contriburion).

Preface

Labels are a basic map component especially with the increase use of multilingual maps. This Chapter is an attempt to provide empirical evidence of users' preference regarding label design. Four fundamental lettering systems (typographies) were implemented in online tests. The first test included maps with Latin labels which were presented to Dutch speakers, the second test included maps with Cyrillic labels which were presented to Bulgarian speakers, the third test included maps with Arabic labels which were presented to Uyghur speakers, and the final test included maps with Chinese labels which were presented to Chinese speakers. The entire tests had identical maps apart from their labels. For each test, a trail of screen maps was presented to users, where users expressed their preference of labels' design by choosing one of the two presented maps. Labels' design included variations regarding labels' size, shape, orientation, and texture, which were associated with both areal and point data. Results were statistically analysed and comparisons were made. The comparison between the applications of design's parameters on labels suggested the frames in which cartographic designers can use when labelling Latin, Cyrillic, Arabic, and Chinese maps, for both monolingual and multilingual maps. Differences and similarities of users' preference were located between cartographic text designs. To produce user friendly screen maps, these results showed that Latin and Cyrillic designs can use identical parameters, but Latin, Arabic, and Chinese designs differ in regard to users' preference of cartographic text design. The results indicated the need to elaborate on testing users' preference of labels' design parameters, considering the combination of different lettering systems on the same tested maps.

5.1 Introduction

Worldwide, maps are the language of geography, since all the geographical phenomena and features can be represented by maps (topographic, thematic, cadastral…etc.). Nevertheless, the language, which the map is depicted with, is crucial for the map users to accomplish map readability and perception. To get the map message through, maps' typography shall be understandable for map users. Users find reading maps so difficult when they had to use maps with languages that they don't frequently read, or even languages that are different from their mother tongue (Deeb and De Maeyer, 2010). Although map symbology is universal and their design is somewhat standardized over the majority of map types, map designers shall take into account typographic design rules to provide the highest level of legibility (Deeb *et al.,* 2013b). Bertin defined seven visual variable 1967; size, shape, value, colour, orientation, texture, and X,Y position (placement), which were the basic principles of cartographic design and they were employed in many of the cartographic work. The earliest studies of typographic legibility on maps were presented by Bartz (1970a, 1970b). Since then, few studies tackled the usability of typographic designs on maps to provide better legibility. Like Phillips' (1981) who studied the label design in view point of letters' shape and case style, using eye tracking technique. Deeb *et al.* (2011, 2013b) studied users' preference of typographic design taking into account different users' categories and different map types and colours. In addition to that, they studied the efficiency of the typographic design using Bertin's visual variables as a basic of label design (2013a), as they differentiated between two categories of users considering their expertise and gender. All the mentioned studies considered Latin as a language of the map. Nevertheless, it shall not be neglected that many maps are designed with different lettering systems. Some lettering systems are symbolic like Chinese, Japanese, Korean; other systems are composed of letters with spacial characteristic. For example; some systems are written from write to left (Arabic and Hebrew) and many others are written from left to right. However, this study handles four lettering systems in two stages. The first stage is a comparison between two of the most used lettering systems, which are Latin and Cyrillic. The second stage is a comparison between three distinct lettering systems, which are Latin, Arabic, and Chinese. Lettering systems use within maps is studied preference wise. Therefore, four parallel experiments were presented to map users to test their preference towards the typographic design as it will be explained in the next section.

5.2 Study design

5.2.1 Apparatus

Four parallel controlled experiments were on line conducted. The first experiment was conducted in the laboratory of the department of geography, Ghent University, Ghent, Belgium. The second experiment was conducted in the department of photogrammetry and cartography laboratory, university of Architecture, civil engineering and geodesy, Sofia, Bulgaria. Both the third and fourth experiments were conducted in the laboratories of Xinjiang institute of ecology and geography, Chinese academy of sciences, China. The stimuli were displayed on 17-inch screens (1280×1024 pixels; refresh rate: 60 Hz). The first experiment is called The Latin test, since maps' typography was depicted in Latin letters and the interviewees were Dutch speakers, whose mother tongue is Dutch. The second experiment is called the Cyrillic test, where maps' typography was depicted in Cyrillic and the interviewees were Bulgarian speakers, who speak Bulgarian as a mother tongue. Whereas the third experiment is called the Arabic test in which the cartographic text was scripted in Arabic and presented to Uyghur interviewees whose language is written in Arabic lettering system. Lastly, the fourth experiment is called the Chinese test as maps' typography was Chinese scripted and the test was presented to Chinese map users.

5.2.2 Participants

Four groups of users participated in the study. First group followed the Latin test. This group consisted of 50 homogeneous Dutch speakers; 25 experts who were in daily contact with maps and they had obtained a Master degree of geographical sciences, and 25 novices who had just started their cartographic education and thus they had no previous cartographic training. The second group followed the Cyrillic test. It is consisted of 50 Bulgarian speakers. Similarly to the Latin test group, the group consisted of 25 experts and 25 novices. The Third group, who followed the Arabic test, consisted of 25 experts and 25 novices Uyghur speakers, who write and read using trained to read Arabic lettering system. In consistence with the previous three groups, the fourth group consisted of 25 expert and 25 novices Chinese who were appointed to follow the Chinese test. In total, 200 participants contributed to this research.

Table 5.1: Bertin's variables on text

Variable		Latin	Cyrillic
Size	Normal	Cartography Cartography Cartography	Картография Картография Картография
	Bold	**Cartography** **Cartography** **Cartography**	**Картография** **Картография** **Картография**
Shape		Cartography Cartography	Картография Картография
Orientation *Within the word*		Cartography *Cartography*	Картография *Картография*
Orientation *Of the word*		**Cartography** **Cartography**	**Картография** **Картография**
Texture		Cartography Cartography Cartography	Картография Картография Картография

5.2.3 Maps and typographic design

Bertin's visual variables were used as parameters to change the typographic design. Latin, Cyrillic, Arabic, and Chinese typography used the same variables simultaneously. Four of Bertins' variables were applied in this study, which are size, shape, orientation, and texture. Four consecutive sizes (8, 10, 12, and 14) were applied to maps' labels. In addition to that, bold sizes were examined (8 bold, 10 bold, 12 bold, and 14 bold). Typographic shape is linked directly to font types, which are infinitive, nevertheless, both Latin and Cyrillic can be classified as serif and sans serif. This classification was extended to cover both Arabic and Chinese lettering systems. The orientation of labels was studied in two levels. Firstly, the orientation of the whole word in relation to shapes of the labelled areal features. Secondly, the orientation of the letters within the words (Italic versus normal). The hierarchy of labels is represented by Labels' texture Table 5.1 shows examples of how the visual variables were implemented in the design (Latin and Cyrillic comparisons). Both areal and point data were depicted in the test's maps (Latin and Cyrillic comparisons). When the two basic map designs were presented to users, labels' placements were considered and both the typography and the letter system (for the four lettering systems) have changed along the maps. Toponyms were designed carefully providing two conditions; firstly, the length of the names, which is controlled by the number of the used letters, were considered (apart from the Chinese test), and secondly, all the toponyms were fictive to prevent any influence of users' previous knowledge on the collected results.

5.2.4 Stimuli, task and procedure

Using Lime-survey, screen maps were presented to users. Four identical stimuli designs were presented to the interviewee, each of which had a sole difference of the typographic system design. The first stimuli introduced maps with Latin lettering design, the second one introduced maps with Cyrillic lettering design, the third presented maps with Arabic lettering system, and finally, the Chinese stimuli presented maps with Chinese lettering systems. Two or three maps were presented to the participant simultaneously for each view during the test trail. Interviewees had to choose one of the presented maps as their preferred map and thusthey would have chosen their preferred typographic design. Once the interviewee chose his/her preferred map, the record will be registered in the database. The sum of records for each variable is considered as a weight which will represent the strength of this variable in the analysis. A pre-test questioner was presented to all participants before they started the test. The questionnaire investigate participants' characteristic such as age, gender, level of expertise, language background, the frequency of using maps, and education. The four stimuli were divided into two stages of analysis. First, lettering systems with similar characteristics were compared (Latin versus Cyrillic). Second, three divers lettering systems were compared (Latin, Arabic, and Chinese).

5.2.5 Data and recordings

Four datasets resulted from the four stimuli. Each of which relates to its test (Latin, Cyrillic, Arabic, and Chinese). All records were online collected and automatically stored in a data base. The records were classified into groups coherently with their represented visual variable. From each data set the parallel records were compared and statistically analyzed.

5.3 Results

As mentioned earlier, the study was design to make its analysis in two stages in which the first stage will compare between two lettering system that have similar characteristic. The second stage will consider three lettering system that are different in nature (Latin, Arabic, and Chinese).

5.3.1 Size

5.3.1.1 Latin versus Cyrillic

Arial font was applied on both Latin and Cyrillic lettering system. Maps were presented at the scale of 1 : 10 000. Table 5.2 shows users' preference of normal typographic sizes. The comparison shows that size 14 is the least preferred size in comparison with sizes 8, 10, and 12. This result is valid for both Latin and Cyrillic. Both lettering systems showed similar trend when sizes were pair wais compared. In addition to that, this comparison showed no significant difference between users' preferences (F=10.098, P=0.000).

Table 5.2: Typographic size comparison for both Latin and Cyrillic (vertical sizes compared to horizontal sizes)

Latin	8	10	12	14
8		23	31	35
10	27		39	46
12	19	11		44
14	15	4	6	

Cyrillic	8	10	12	14
8		27	34	42
10	23		40	47
12	16	10		45
14	8	3	5	

Bold sizes were compared pair wise with normal sizes. Figure 5.1 illustrates the trend that users showed over the four point sizes (8, 10, 12, and 14). ANOVA test was applied to compare between the normal and the bold sizes of both Latin and Cyrillic and it showed no significant difference between their trend (F=0.004, P=0.954). For bold design, on the one hand, size 8 is the most preferred and size 14 is the least preferred. A small fluctuation in the Cyrillic trend appeared between size 10 and size 12. On the other hand, the vice versa trend occurred for normal design.

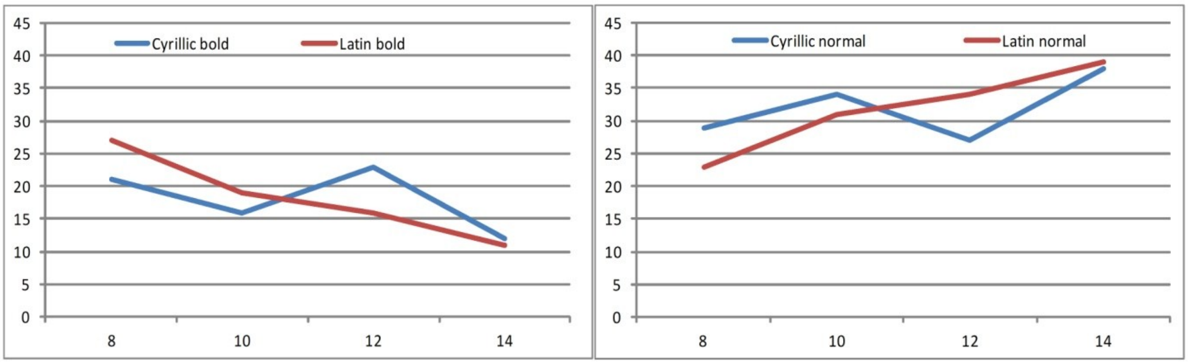


Figure 5.1: Both bold and normal size comparison for Cyrillic and Latin lettering systems

5.3.1.2 Latin, Arabic, and Chinese

Four consecutive sizes (8, 10, 12, and 14) were depicted in Arial font for Latin, Tahoma font for Arabic, while Chinese were depicted in (Microsoft JhengHei). To test users' preference regarding the difference between Latin, Arabic, and Chinese lettering systems size wise, ANOVA test was conducted to test the influence of size differences on users' preference when bold sizes were compared to normal sizes. The test showed significant influence of letters' size (normal versus bold) on users' preference (F=10.230, P=0.005). Therefore Tukey post-hoc test was conducted to explore between which lettering systems the difference was. The test showed significant difference between Latin compared to Chinese users' preference (F=4.646, P=0.046) and between Arabic compared to Chinese (F=4.646, P=0.004). Figure 5.2 shows users' preference of normal sizes compared to bold sizes for the three lettering systems.

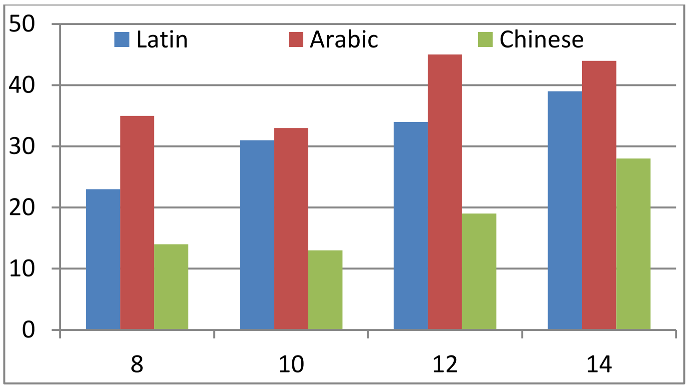


Figure 5.2: Normal size comparison to bold sizes for Latin, Arabic, and Chinese lettering systems

To acquire and overview of the users' responses regarding each lettering system and cartographic text size, a triangulation model was developed. This model combines users' preference of the three lettering systems which were presented as percentage in each triangle line. The subtracted area between the three users' responses (their preference) represent the 'popularity' of each size design (normal versus bold). This model shows that size 8 normal and size 10 normal are almost as proffered as their corresponding bold sizes. But on the contrary, the graph shows that normal 12 and 14 normal sizes should be used instead of their corresponding bold sizes to acquire a combined users' preference. It worth mentioning that size 14 bold cannot acquire a combined users' preference of the three lettering systems at the same time. See Figure 5.3.

5.3.2 Shape

5.3.2.1 Latin versus Cyrillic

Since the shapes of lettering systems are linked directly to letter fonts. The list of letters' shapes falls into different categories; historical, automated, artistic, …etc. Additionally, for both Latin and Cyrillic shapes can be classified as serif and sans serif. Arial font was chosen as representative for sans serif fonts (Figure 5.4) and Times New Roman as a representative for serif font. When comparing users' preference for shape design by using one-sample T test, the results showed no significant difference between both lettering systems (F=0.810, P=0.463).

5.3.2.2 Latin, Arabic, and Chinese

Unlike the situation of comparing the shapes of Latin with Cyrillic, where both serif and sans serif terms are well known for their typography, but using a known term to describe such characteristic in the comparison between Latin, Arabic, and Chinese is not applicable. Therefore we will describe two groups in which the characteristics of both serif and sans serif are attributed. With Group A we refer to the sans serif fonts or fonts that have similar characteristics (the least strokes the possible), and with group B we refer to serif fonts or fonts that have similar attributes (extra strokes included). Thus, A and B fonts were applied on each lettering system using normal size 10. Then, a comparison was made for Arial versus Times New Roman for Latin, Tahoma Versus Times New Roman for Arabic and Microsoft JhengHei versus SimSun for Chinese.

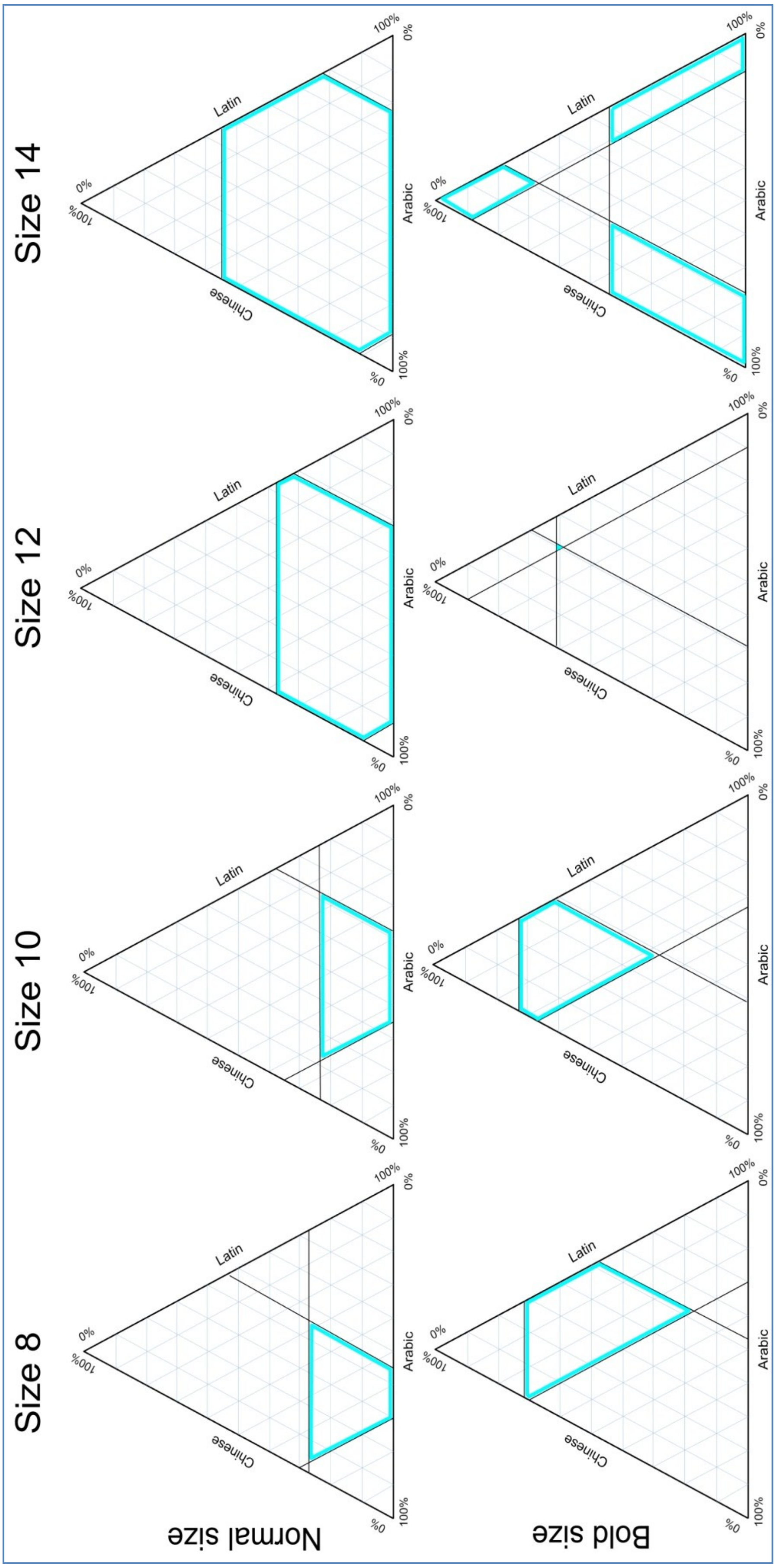


Figure 5.3: Normal size comparison to bold sizes for Latin, Arabic, and Chinese lettering system

Figure 5.4: The application of Arial shape on both Latin and Cyrillic lettering system

In parallel, shape A represents the sans serif and its resemblances (Arial, Tahoma, Microsoft JhengHei), and shape B represents the serif font and its resemblances (Times New Roman, and SimSun). ANOVA test showed that users' preference of shapes was significantly different when comparing between shape A and shape B (F=846.0, P=0.000), as they preferred shape A over shape B. Whereas the difference between users' preference of the three different lettering system was not significantly different (F=0.000, P=1.000).

5.3.3 Orientation

The orientation of labels was studied in two phases. The first phase tackled the overall orientation of label in relation to the depicted feature. The second phase tackled the orientation of the letters within the world (italic versus normal). Figure 5.5 shows examples of the overall label orientation.

5.3.3.1 Latin versus Cyrillic

Three basic orientation of the overall orientation were studied; horizontal label orientation when all labels were placed in horizontal manner, tilted orientation when labels were placed in regard to the depicted features' shape, and finally mixed orientation which combines the two previously mentioned orientation. See Figure 5.5. The orientation of labels showed no significant difference between Latin and Cyrillic (F=0.005, P=0.945). Similar trend of users' preference for both lettering system, occurred and illustrated in Figure 5.6.

When italic labels were compared with straight labels (upright labels), four sizes were engaged in this comparison (8, 10, 12, and 14). Straight label designs were preferred over italic label design along the four tested sizes. The comparison between the lettering systems showed no significant difference between Latin and Cyrillic (F=4.528, P=0.077).

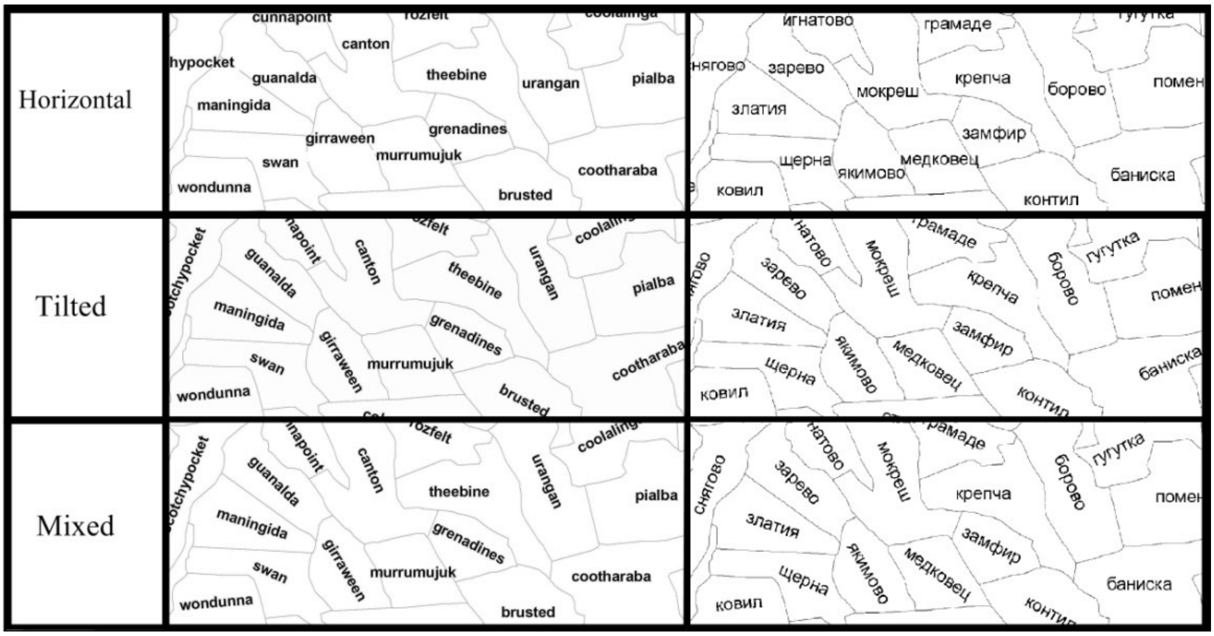


Figure 5.5: The application of the overall label orientation on both Latin and Cyrillic lettering system

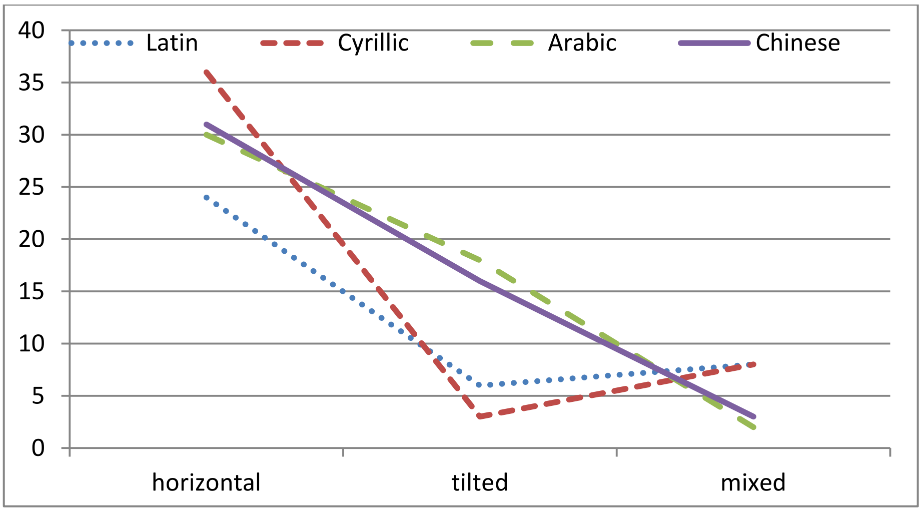


Figure 5.6: Users' preference of Label orientation for the entire lettering system

5.3.3.2 Latin, Arabic, and Chinese

Represented in Figure 5.6, the trend of users' preference of cartographic text orientation for Latin, Arabic, and Chinese. The orientation of labels showed no significant difference between Latin, Arabic, and Chinese (F=0.098, P=0.908), when labels were placed horizontally, tilted according to the shape of the labelled area, and finally mixed placement, which combines the two previously mention orientations.

When italic labels were compared with straight labels, four sizes were engaged in this comparison (8, 10, 12, and 14). The comparison between straight label design and italic label design along the four tested sizes showed no significant difference between both designs (F=0.173, P=0.911) size wise, when ANOVA test was used. The comparison between the lettering systems using ANOVA test showed a significant difference of using different lettering systems (F=8.945, P=0.007)

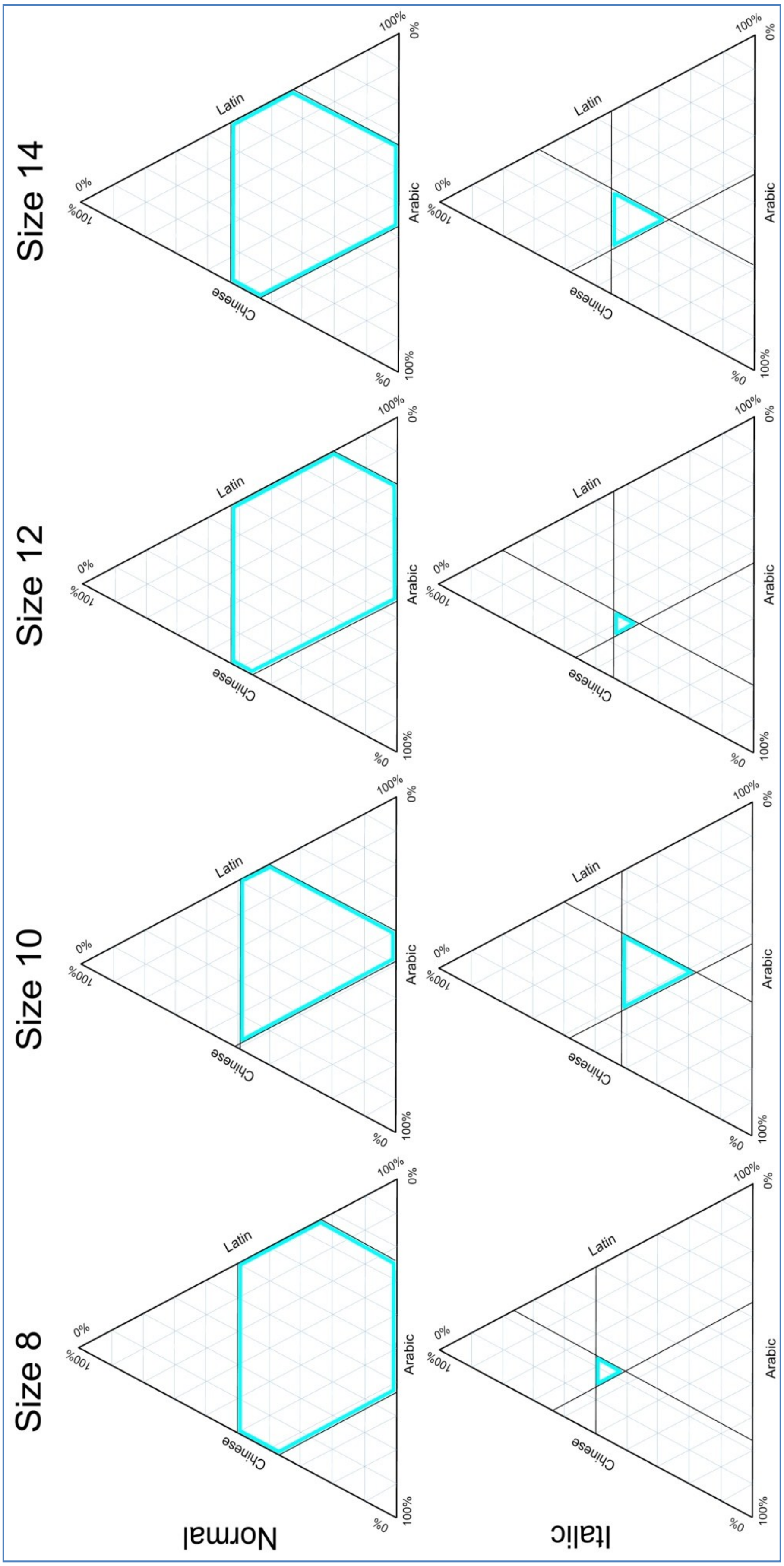


Figure 5.7: Normal and italic comparison for Latin, Arabic, and Chinese lettering systems

The post-hoc Tukey test followed to indicate where the difference was between the three lettering systems. The test showed that users' preference of Latin italics' was significantly different from users' preference of both Arabic Italic (F=2.131, P=0.040) and Chinese italic (F=2.131, P=0.007).

To illustrate how users' preference of the three lettering systems can be combined, the triangulation model was applied to present the percentage of users' preference of each lettering systems. The subtracted areas represent the popularity of each design regarding the combination of the three lettering systems. The results show that normal (upright) orientation is more preferred than italic orientation. These results are valid for the four tested sizes. See Figure 5.7.

5.3.4 Texture

Texture was formed by the combination of two variables at least. Hierarchal texture involves different typographic sizes. To form texture, a comparison between two shapes (serif versus sans serif within hierarchal texture) and weight size (bold versus normal within hierarchal texture) was made.

5.3.4.1 Latin versus Cyrillic

The first tested textural design was formed by the combination of three hierarchal normal sizes (8, 10, and 12) and two shapes (Arial versus Times New Roman). The second tested texture design was formed by the combination of three hierarchal bold sizes (8, 10, and 12) and the size weight (normal versus bold).

The first textural comparison (illustrated in Figure 5.8) describes users' preference when sans serif font (Arial) and serif font (Times New Roman) were presented in Hieratical way. This comparison was made in two phase; first normal sizes were compared, and then bold size was added to the comparison. Similar trend appeared in the comparison between Arial and Times New Roman. Texture with sans serif font is more preferred than texture with serif font for both Latin and Cyrillic and in both normal and bold design.

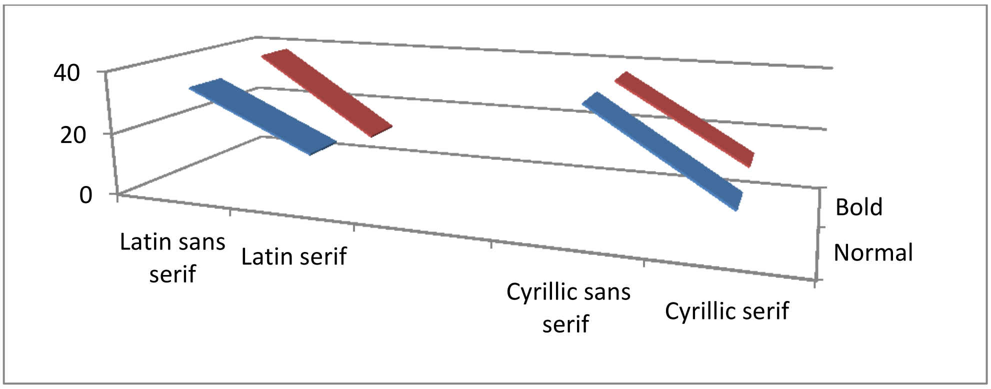


Figure 5.8: Users' preference of Labels' texture (sans serif versus serif) for both Latin and Cyrillic

The second textural comparison (illustrated in Figure 5.9) describes users' preference when normal size and bold size were presented in hierarchal way. This comparison was also made in two phase; first normal sizes were compared with bold sizes for Arial font, and then normal sizes were compared with bold sizes for Times New Roman font. On the one hand, Texture of Cyrillic showed similar trend for both sans serif and serif font as users preferred texture of normal sizes over texture with bold sizes. On the other hand, Texture of Latin did not show similar trend. Users' preference of both normal and bold size was almost the same with sans serif. But with serif font users preference of hierarchal texture was higher for bold size.

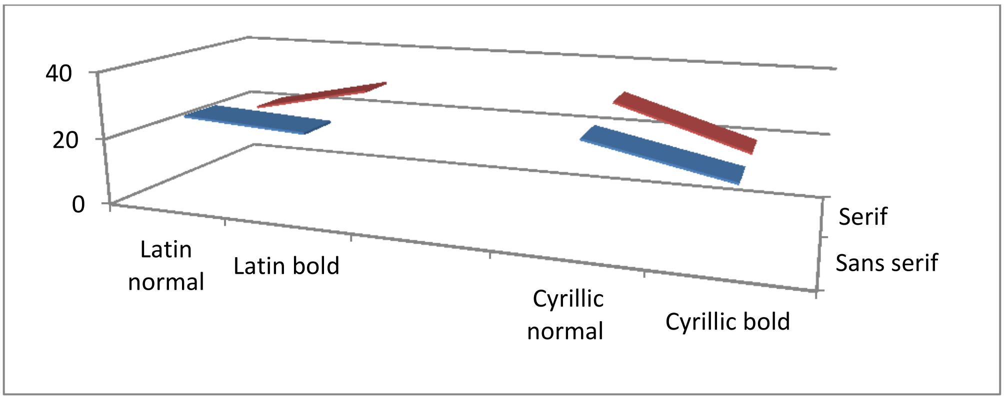


Figure 5.9: Users' preference of Labels' texture (normal versus bold) for both Latin and Cyrillic

5.3.4.2 Latin, Arabic, and Chinese

Labels' texture will change when the shape of labels changes and when their sizes change as well. The hierarchy representations were studied through the Texture variability, three consecutive size (8, 10, and 12) were used and two different fonts resembling the serif (Shape A) and sans serif (Shape B) for the three lettering systems (as explained earlier). Normal Hierarchy was compared to bold hierarchy for both shape A and shape B, illustrated in Figure 5.10. When comparing between users' preference of the presented textural designs, in regard to the three lettering systems, no significant difference was found between users' preference of Latin, Arabic, and Chinese (F= 0.554, P= 0.624).

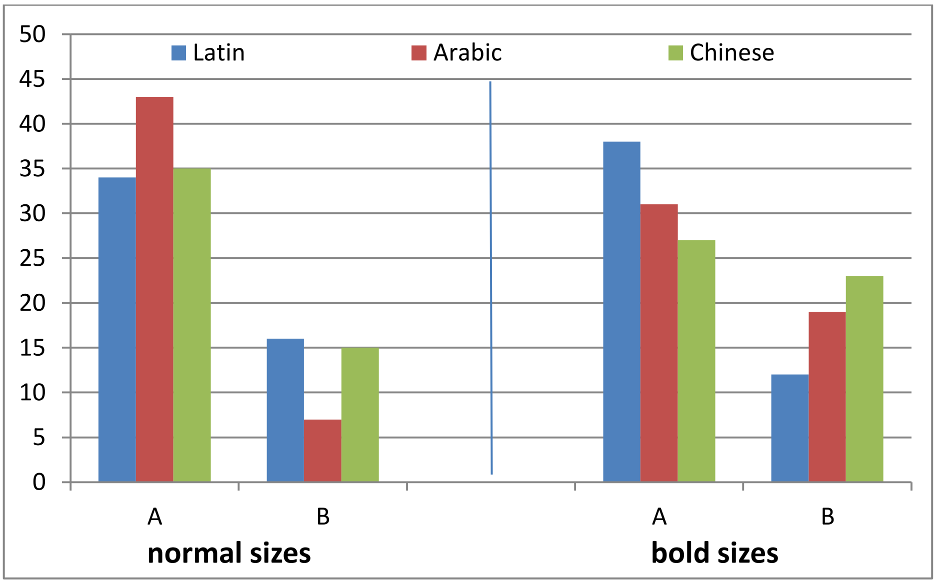


Figure 5.10: Users' preference of Labels' texture (normal versus bold) for Latin, Arabic, and Chinese

5.4 Discussion

As the analysis was made in two stages regarding the lettering systems' characteristics, the presented discussion follows this division.

The research compared in the first stage between tow lettering systems that have similar characteristics. Latin and Cyrillic are similar in the way words are formed, the direction of writing words (from left to right). Fixed shapes of letters in regards to their contextual letters' shapes, and similar use of uppercase styles, italics…etc.

Size, shape, orientation, and texture show similar users preference in this experiment. This result can be linked basically to the similarity of the letters' design themselves in both systems and the way they form words, unlike what we found in the second stage of this experiment.

The second stage of this study showed that Latin, Arabic, and Chinese have a significant influence on users' preference. Visual variables (size, shape, orientation, and texture) applications on each lettering systems resulted in different users' preference. Basically, the significant difference could have resulted because of the differences in lettering systems' characteristics.

Distinguishable properties are included in each lettering system. For example, Latin letters have fixed contextual shape, unlike Arabic whose letters' forms differ according to their location in the word (contextual shape). Meanwhile, the 'letters/words' in Chinese are pictographically formed. In addition to that, the directionality of writing differs as Latin is written from left to right, Arabic is written from right to left, and Chinese could be written either from left to right or from top to bottom. Many more cultural differences between the lettering systems can be identified. These differences are reasonable causes for the significantly different results in the second stage of the test.

During setting the application of Bertin's variable some important issues were noticed:

- although letter size is defined by its corresponding point size, when comparing between five letters toponyms written in the same point size, one is in Arabic and the other is Latin, each consumes different space on the map;
- the relation between typographic size and shape is crucial. At a fixed size and within a specific lettering system, different typographic shapes could appear in different sizes, and thus, they would consume different space on mapface;
- at a fixed size and for the same shape characteristics, different lettering system could appear in different sizes. This issue was noticed when we studied shape A (sans serif and its resemblances). Arial was applied on Latin, Tahoma was applied on Arabic, and Microsoft JhengHei for Chinese;
- Considering shape A and shape B characteristics, their application on Chinese typography did not show noticeable difference when they were depicted in the smallest sizes (8 and 10).

The triangulation model was a good tool to visualize and asses the combination of users' preference regarding the three lettering system. Drawing the individual lines that represent the percentage of the collected data, resulted in a polygon, which may differ in size according to the positivity of the combination of the three variables (lettering system in our study). The more positive the combination, the larger the polygon is. In the case where the combination of the responses is negative, three disconnected polygons will result, which states that such combination cannot be done (the three lettering system in our study, Figure 5.3, the last triangle).

5.5 Conclusion and Future work

Users' preference of the visual variables' application on both Latin and Cyrillic typography showed similar trend of most variables. In addition to that, no significant difference was found between Latin and Cyrillic design. Visual variables can be applied similarly and in parallel for both lettering systems. Regarding Latin, Arabic, and Chinese lettering system, some significant differences occurred when applying the visual variables regarding users' preference. This issue shows that lettering systems' characteristics has an influence on users' preference.

As we move towards multilingualism, which become a widespread phenomenon in the economically globalized word, interactions between cultures occur daily. It is important to consider multilingual mapping in future work where further experimental conditions shall be considered, especially the combination of different lettering systems on map to define clearly the conditions of using visual variables for multilingual maping.

5.6 References

Bertin, J. 1967. Semiology of Graphics: Diagram, Networks, Maps, Translated by William J. Berg, 2011, California: Esri press.

Bartz, B.S. 1970b. Experimental Use of The Search Task in The Analysis of Type Legibility in Cartography, The Cartographic Journal, 7(2), pp.103-113.

Bartz, B.S.1970a. An Analysis of the Typographic Legibility Literature, Assessment of Its Applicability to Cartography, The Cartographic Journal, 7(1), pp.6-17.

Deeb, R., Ooms, K., Vanopbroeke, V., & De Maeyer, P. 2014. Evaluating the Efficiency of Typographic Design: Gender and Expertise Variation, The Cartographic Journal, 51(1), pp.75-86.

Deeb, R., Ooms, K., Van etvelde, V., & De Maeyer, P. In press. Toward a deeper understanding of cartographic text visualisation: assessment of user preferences and colour influence. The Cartographic Journal, DOI:http://dx.doi.org/10.1179/1743277413Y.0000000073

Deeb, R., Ooms, K. & De Maeyer, P. 2012.Typography in the Eyes of Bertin, Gender and Expertise Variation, The Cartographic Journal, 49(2), pp.176-185.

Deeb, R. & De Maeyer, P. 2010. An Interprative Phenomenological Analysis of The Noise Occuring in Reading Maps.The workshope on IPA methodology, Doctoral School, Ghent University.

Phillips, R.J., 1981. Search for Target in a Random Arrangement of Names: an Eye Fixation Analysis, Canadian Journal of Psychology, 35(4), pp.330-346.

6

Cartographic Text, The Efficiency of Label Design, and Cross-Cultural Comparisons

Modified from: Deeb,R., Ooms, K., Kurban,A., Ablikim, A., & De Maeyer, P.2015. On the Cartographic Text Design: Users' Efficiency and Cross Culture Comparisons, under review, Information Visualization.

Preface

Maps are exceptionally sophisticated visual displays in which text is distributed irregularly, often using a large number of styles, and the demands on the reader are very different from other types of display. An optimal design of the labels improves their usability, particularly the efficiency, which allows users to interpret the map contents more efficiently. However, the users' efficiency regarding cartographic text is linked not only to design and placement aspects but also to the language characteristics of the text. This chapter presents a user study that compares maps using three different lettering systems (Latin, Arabic, and Chinese) in a visual search task. Three sets of stimuli were designed for native users who were asked to locate a target label on the map. In the set of stimuli, variations in the labels' design were introduced (i.e., size, shape, orientation, and texture). Reaction time measurements (indicating when a target label was found) were registered to measure the participants' efficiency. The within lettering systems analysis showed significant differences over size and orientation, as determined via either a one-sample t test or one-way ANOVA for each of the studied visual variables. However, the between lettering systems designs analysis was not significantly different from orientation and texture, as determined by ANOVA. The requirements of efficient label design were described for both the within and between lettering system efficiency. Consequently, label size, shape, orientation and texture design can be integrated to obtain more efficient maps for both monolingual maps and multilingual maps.

6.1 Introduction

Multilingualism is a widespread phenomenon, as there are more than 6,800 languages across the globe. Various lettering systems exist around the world due to language, cultural and religious differences. Noticeably, we are moving into more interactions between cultures. Businesses, governments, and nongovernmental organizations bring goods and services to every facet of society in every part of the earth. To obtain location-based information about places of interest, various geospatial products with different lettering systems are in the world's service.

Currently, one of the most important visual communication products is the map, especially when a spatial component is involved. To fully communicate with the audience, the cartographer employs symbols and words. While some cartographers consider lettering to be a standard cartographic element (Gerber, 1981), others consider it to be the fourth symbol type in addition to point, line, and area (Fairbairn, 1993). In both cases, the syntax of letterings does not resemble the design of other symbol types because it is directly linked to a language and its characteristics.

Researchers have thoroughly discussed the efficiency of Latin typographic design on various scripts' layouts to allocate better legibility of different designs. On plain text, Bernard *et al.* (2003) discussed the legibility of Times New Roman and Arial scripts on computer displays, taking into account the labels' shape and size. They found that 12 point size was more efficient (faster to read) than 10 point size. In addition, they found that Times New Roman is less efficient than Arial. Arditi and Cho (2005) tested the legibility of serif and sans serif fonts and contradicted Bernard *et al.* when they found that serif font was slightly more legible than sans serif font. In addition, they discussed the influence of the characters' case style on text legibility (Arditi and Cho, 2005), and they linked the efficiency of upper case style to letters' size in comparison with the letter size of lower case style.

Although several studies were presented to address the legibility of typography on plain text, cartographers discussed the use of these findings to improve the legibility of maps. Some of them invited us to use these findings as a basis for cartographic text design (Bartz, 1970a), but others argued that typographical guidelines defined from experimental work in other fields cannot be safely applied to maps (Taylor and Hopkin, 1975; Phillips *et al.*, 1977; Phillips *et al.*, 1978; Deeb *et al.*, 2012,2013). Following the later group of researchers, the study aims to examine label designs in different lettering systems. As the chapter presents an empirical evidence of how cartographic text designs can improve its legibility.

6.2 Characteristics of different lettering systems

Because the characteristics of lettering systems also have a vital impact on their design and thus on users' efficiency towards these designs, they are implemented in this study in which the Latin, Arabic and Chinese system will be included. These three lettering systems are very distinct in a number of ways. For example, the Latin lettering system has 26 characters with some diacritic marks (varying in number according to the language), the Arabic lettering system has 28 characters with eight basic diacritic marks, and the Chinese lettering system has 70244 characters (included in GB 18030-2005, the Chinese national standards issued by the Standardization Administration of China, SAC) and some pictographic diacritic marks, but only approximately 3500 characters are commonly used as pictographic letters. The number of characters is not the only variation among the three lettering systems. However, the most significant variation is the way in which the characters are gathered to form a word: the Latin words are formed by gathering the individual letters from left to right in their fixed shape, both in upper case and/or lower case style; the Arabic words are formed by gathering the letters from right to left in their contextual shape, where the letters' forms change, depending on their position in the word (isolated, initial, middle or final); the Chinese characters form pictographic words, as one character can stand for a whole word (meaning forming a one-syllable word) or a single-syllable part of a word. Chinese words can be written from left to right (normally) or from top to bottom. Figure 6.1 illustrates the contextual shape of Arabic letters.

shape	**Initial**	**middle**	**Final**	**pronunciation**
ج	جـ	ـجـ	ـج	J
م	مـ	ـمـ	ـم	M
ك	كـ	ـكـ	ـك	K
ب	بـ	ـبـ	ـب	B
ر	ر	ـر	ـر	R
ه	هـ	ـهـ	ـه	H

Figure 6.1: The contextual shape of Arabic letters (isolated, initial, middle, and final)

The efficiency of the design of cartographic texts in the Latin lettering system has already been the subject of a number of user studies in the past. Initially, it was studied by Bartz (1970 b), who investigated users' efficiency on paper maps using three label variations: shape (font), size and hierarchy (combination of different case styles). She located variations in the registered search times for each label variation, especially when users were informed about the target labels' variation. Phillips (1981) tested the influence of other criteria with regards to word length, the word's initial letter and its overall shape (cap height, x- height, stem, bowl, etc.) and concluded that these elements significantly affect the search time. In addition, sans serif fonts over the map face were tested by Feldmannand and Kreiter (2006) on topographic maps. The controlled test included Arial, Univers and Frutiger fonts and resulted in similar efficiency relative to the sans serif fonts. The above-mentioned studies identified some outlines for efficient label design, provided that maps are presented on papers. Later, van den Worm (2001) emphasized that the readability of labels is expected to be affected by the typographic design. He suggested using bold typography for better readability for web maps. In addition, he recommended using sizes larger than 10 point when rasterizing the map as a general rule. Complex criteria of the cartographic text design were linked with Bertin's (1967) visual variables by Deeb *et al.* (2013), who described how size, shape, orientation, colour, value and texture can functionally be used in label design. They analysed the efficiency of different map users, taking into account their characteristics and training background. Their results defined which visual variables to use (on cartographic text) for certain purposes. Although their research used the basic visual design principles, their findings are limited to the Latin lettering system because of the special characteristics embedded within the lettering system.

Fundamental differences from the Latin lettering system can be found in the Arabic lettering system, where some issues conflict with their design, including the contextual shape of letters, variations in contextual letter sizes, rasterizing problems and writing direction. The contextual shape is a critical legibility issue that includes variations in cursivity. This cursivity implies four different styles of the same letter, which are based on their location in the word: initial, middle, final or separated (see Figure 6.1). Additionally, the Arabic lettering system implies ligatures; i.e., the way in which the letters are connected, where some ligatures are mandatory and others are optional, existing only for aesthetic reasons, legibility, or justification; see Figure 6.2. Furthermore, allographs involve a change in a letter's shape with respect to the neighbouring letters' positions and shapes (the contextual shape is a simple example). Moreover, the kashida rules, i.e., the connection distance between letters (Azmi and Alsaiari, 2010), provide emphasis, legibility, aesthetics and justification. These elements introduce only small conflicts for the Arabic typographic design. More critical issues arise when using the Arabic lettering system in cartographic text design. Because the various attempts to solve the problem of map lettering are closely related to the tools and media available (Woodward, 1987), rasterizing the maps and the resulting low-resolution products highly affect cartographic text legibility (Trenkle *et al.*, 2001; Al-Harkan and Ramadan, 2005). Although rasterizing a map avoids text changes at the users' side, rasterization decreases the map's legibility. Therefore, some treatments are conducted to secure the transformation from vector format to raster format, such as anti-aliasing (van den Worm, 2001) and hinting (Hersch, 1994).

Enlarging the Arabic letter size can increase the text's legibility on a map, but it can also cause the text to cover a considerable portion of map features and might not be compatible with the label's function. Therefore, it is crucial to learn how the visual variables are applied onto the cartographic text and what function they can serve, such as supporting efficient map reading, providing the primary function of addressing the geospatial data and providing the secondary function of indicating the nature of the represented object (Kraak and Ormeling, 2010).

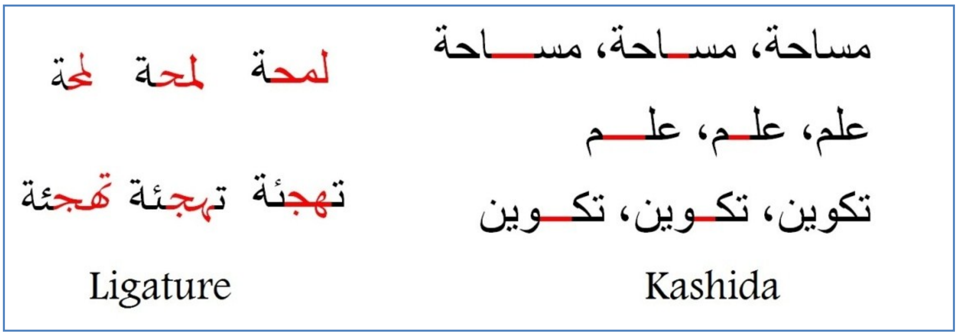


Figure 6.2: The differences made on the word caused by ligature (left) and kashida (right)

Remarkably different from both Latin and Arabic lettering systems, the Chinese lettering system is composed of many pictographic symbols. These symbols are drawn using strokes that fall into eight main categories (horizontal, vertical, left-falling, right-falling, rising, dot, hook and turning). For plain text, the efficiency of Chinese symbol size was studied based on point measurements, character heights and display resolutions (Chan and Lee, 2005; Huang *et al.*, 2009), where they concluded that a higher resolution combined with a larger font size and more line spacing provides faster reading times and higher preferences than using lower resolution, smaller font sizes and less line spacing. The shape of the Chinese lettering system was also discussed by Chan and Lee (2005), who defined two common font types as serif (Ming) and sans serif (Li) fonts and recommended some settings for better legibility regarding both cases: for label features, the pictorial symbols should be drawn in limited spaces; many conflicts could be introduced when reading these symbols over the map face, especially for smaller sizes, certain fonts and smaller display media.

Earlier studies of the three lettering systems indicated that different designs of text led to different users' responses for text shape (font) and text size, but few of them considered text design in the cartographic context, as most of them did not consider all of the variation that can be applied to text, disregarding text orientation, text colour and value, and text texture over the map face.

Additionally, lettering systems were studied individually. Although the differences of design variability were studied within each lettering system, the connection of typographic variations between the lettering systems was not made. It is worth mentioning that the use of multilingual mapping is a growing concern because many online maps are presented in more than one language; i.e., Google Maps combines, in addition to Latin, other lettering systems in many parts of the globe, such as Chinese, Cyrillic, and Arabic; when the political boundaries gather more than one culture, as in Belgium, labels (Latin) written in Flemish, French and German can be found on one map; in Urumqi, the Latin, Arabic and Chinese lettering systems are presented on one map to meet all audiences' needs.

This chapter presents a combined between- and within-user study in which the three discussed lettering systems were examined in terms of label design and how the visual variables can be used to change the functional designs of labels. The goal of this study is twofold: first, to provide efficient label designs of each lettering system; second, to assess the influence of lettering systems on label design regarding users' efficiency. Therefore, different label designs were implemented in three parallel tests, as explained in the next section.

6.3 Study design

6.3.1 Participants

Three corresponding stimuli were presented to three parallel groups of participants. The first group consisted of 50 Latin native readers who use the Dutch language on a daily basis. Of the 50 Latin reader participants, there were 25 experts who are in daily contact with cartographic products and have acquired at least a master's degree in geography or geomatics. Their average age was 29.5 years. The other 25 participants were novices who did not acquire any type of cartographic training beforehand. The novices were pupils at the secondary school level, with an average age of 16.4 years. The Latin group consisted of 25 females and 25 males, both groups with an average age of 23 years.

The second group consisted of 50 Arabic readers. This group uses the Uighur language, which is written in the Arabic alphabet. This group had 25 experts who obtained a degree in geographic information science and were involved in geographical projects and studies. Their average age was 26.9 years. The others were 25 novices who were not cartographically trained. Their average age was 25.3 years. The Arabic reader novices were bachelor's students specializing in statistics, electronic information and control engineering. The 50 Arabic readers consisted of 27 females with an average age of 24.8 years and 23 males with an average age of 27.6 years.

The third group consisted of 50 native Chinese readers. This group contained 25 experts who received cartographic training and geographic academic education and had an average age of 23.9 years. In addition to the 25 experts, there were 25 novices who were educated in statistics, biology, ecology and electronic information. The novices' average age was 23.5 years. Of the 50 Chinese readers, there were 25 females, with an average age of 22.7 years, and 25 males, with an average age of 24.8 years. All of the 150 participants had a correct or corrected vision, and none of them was colour blind.

6.3.2 Stimuli

Two major map designs were utilized in this user study (see Figure 6.3, left versus right), which were presented on a screen to the users.

The first design represented a thematic map of areal data with a very simple background: i.e., a boundary representation without any background colour. This controlled design was used to limit influencing factors on users' responses (e.g., colours, complex map objects, etc.). The second map design represented a topographical map populated with point labels. This design was used to test the functional aspect of label hierarchies. This topographical map was designed at a scale of 1 : 10 000 and presented at a scale of 1 : 3 700 to the participants. For both maps, fictive and unknown name labels were used to eliminate any bias caused by the users' previous knowledge or specific training for the studied region.

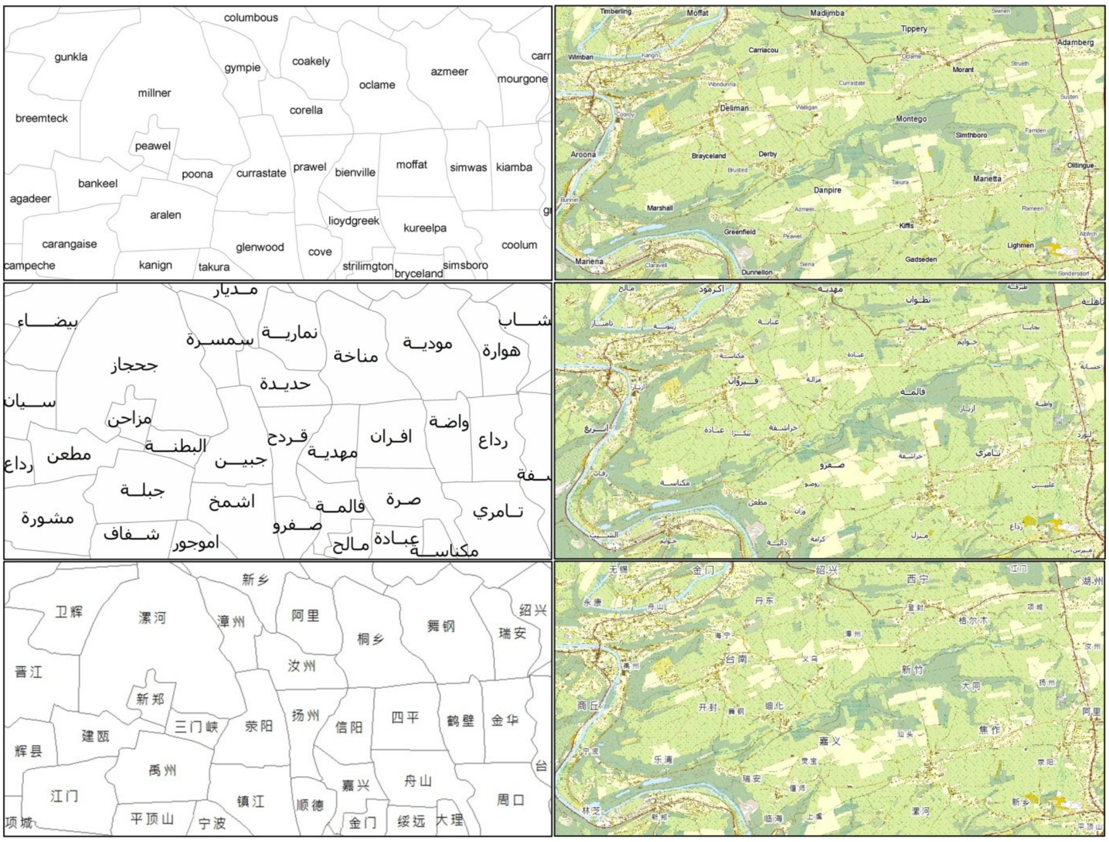


Figure 6.3: Example of the parallel map design for the two map (left versus right) types and three lettering systems (top: Latin; middle: Arabic; bottom: Chinese)

Three corresponding label designs were integrated in the three sets of stimuli (Latin, Arabic and Chinese). The labels were designed with variations in size, shape, orientation and texture. These were applied on the labels individually or in combination with each other. The visual variables were extensively explained by Bertin16, and their application on map labels was justified by Deeb *et al.* (2013). Figure 6.3 also illustrates the three lettering systems that were used for both map designs. The dot was used as a unit to indicate typographic size for the three lettering systems. Four consecutive label sizes were applied - 8, 10, 12 and 14 - on each of the corresponding fonts in the three lettering systems: Arial font for Latin, Tahoma font for Arabic and Microsoft JhengHei for Chinese. In addition, shape was studied by comparing two groups of fonts: first, sans serif font and its resemblance (Arial for Latin, Tahoma for Arabic and Microsoft JhengHei for Chinese); second, serif font and its resemblance (Times New Roman for Latin and Arabic, and SimSun for Chinese). Additionally, orientation was studied in two phases: the overall orientation of the label and the orientation of the individual characters (italic versus normal fonts). Finally, the labels' texture was made by the functional combination of size and shapes, which will be explained in further sections.

A set of 40 thematic and topographic maps was presented to each participant of the three user groups (16 thematic maps and 24 topographic maps). The same order of maps was used for each lettering system, where each participant followed the same order of conditions (map type, typographic variability and target label placement); see annex 1. The design of the thematic maps was identical for all stimuli, apart from the labels, which changed according to the studied typographic variations. Similarly, the same topographic maps were presented to all participants. The topographic maps depict both urban and rural regions. The topographic map sheets were selected out of the Belgian topographic map series on 1 : 10 000, with map sheet 46/8-S for the urban region and map sheet 49/2-N for the rural region. The two maps were used to help separate between label designs that belong to the same variables to reduce the learning process to its lowest, which might have evolved while users were conducting the search task. However, only results from rural maps were analysed (mostly green background) because the background colour does not influence the reading of black labels (Deeb *et al.*, 2014). The order of the experiment conditions is explained thoroughly in appendix 1.

6.3.3 Task

To determine the users' efficiency towards different label designs, a search task was presented. This task was deemed to be the most appropriate task, and it reflects how the design can affect users' effeciency (Bartz, 1970a, phillips *et al.*, 1978). It seems to be a good task for evaluative purposes because it involves both the peripheral and the central vision, whereas tasks such as normal reading or making choices put much less demand on peripheral vision (phillips *et al.*, 1978).

An explanation of the test procedure was presented on the first screen of the test. The participants were informed that they had to execute a visual search task: they had to locate a target label – which was presented (centrally) above the map – in the map image (see Figure 6.4). The target labels (on top) were presented in a neutral font for the three lettering systems (OCR A Extended for Latin, Courier New for Arabic and SimHei for Chinese). Once users found the target label in the map image, they had to click on it; thus, their reaction time was registered (time interval, in milliseconds, between the display of the map and the mouse click). After this action, a new map with a new target label was displayed. The same label was never requested more than twice, and the studied variables were presented such that maps related to the same visual variable were not presented directly after each other. To assure these criteria, the maps were presented in the same order to all participants.

After completing the visual search task on the 40 stimuli, the participants had to complete a post-test questionnaire to provide their gender, age, level of expertise, the most frequently used map types, frequency of map use and other personal characteristics. All participants finished the set of 40 stimuli in 10-15 minutes.

6.3.4 Procedures and apparatus

Three identical tests were designed for the three lettering systems (Latin, Arabic and Chinese), each of which was presented to the corresponding language speakers. Because the users' reaction times were measured, the loading times of all maps needed to be similar. The latency of map loading time was controlled by their size: each map had a similar size (approximately 350 KB) to ensure equal loading times. In addition, the test was run on similar hardware and presented in the same browser (Google Chrome). Furthermore, biases in the measurement due to resolution and size differences were avoided, as all participants completed their test on a 17-inch flat screen with a 1280×1024 resolution. The three controlled tests were implemented in an online questionnaire, the results of which were stored in a database.

Figure 6.4: Example of the test stimuli with the Arabic lettering system

6.3.5 Recordings

Although all participants completed the questionnaire in a controlled environment, some outliers in participants' reaction time were observed. The outliers were excluded from the database because they cannot represent factual reaction times. In the three datasets, 58 outliers of more than [M+2 SD] out of 6000 measurements were located (< 1%). Specifically, because no outliers less than [M-2 SD] were found, these outliers could have resulted from unexpected network problems or from a diminished level of attention by some participants while performing their search. In addition, these outliers were located in the three datasets (Latin, Arabic and Chinese) and associated with different users or variations in the labels' design. Consequently, these values are considered as errors in the complete dataset and could thus negatively contribute to the goal of our research. As a consequence, it was decided to remove all of these outliers from the dataset.

6.4 Results

As described in the stimulus design, four visual variables were implemented in the label designs (size, shape, orientation and texture) and combined with the three lettering systems (Latin, Arabic and Chinese). A between lettering systems analysis was also executed to demonstrate the influence of the lettering system characteristics on the use of the visual variables and, thus, their efficiency. The results and comparisons between the different label design options are discussed in the following sections.

6.4.1 Size

To examine the efficiency of size differences within and between the typographic systems, four consecutive normal (not bold) sizes (8, 10, 12 and 14), measured in dots, were embedded in the analysis of Latin, Arabic and Chinese typography. Additionally, the type sizes were enlarged by making them bold. The resulting normal and bold sizes were studied and compared. Figure 6. 5 shows examples of the Chinese four sizes in both normal and bold representations. The examples are sketches from the test maps.

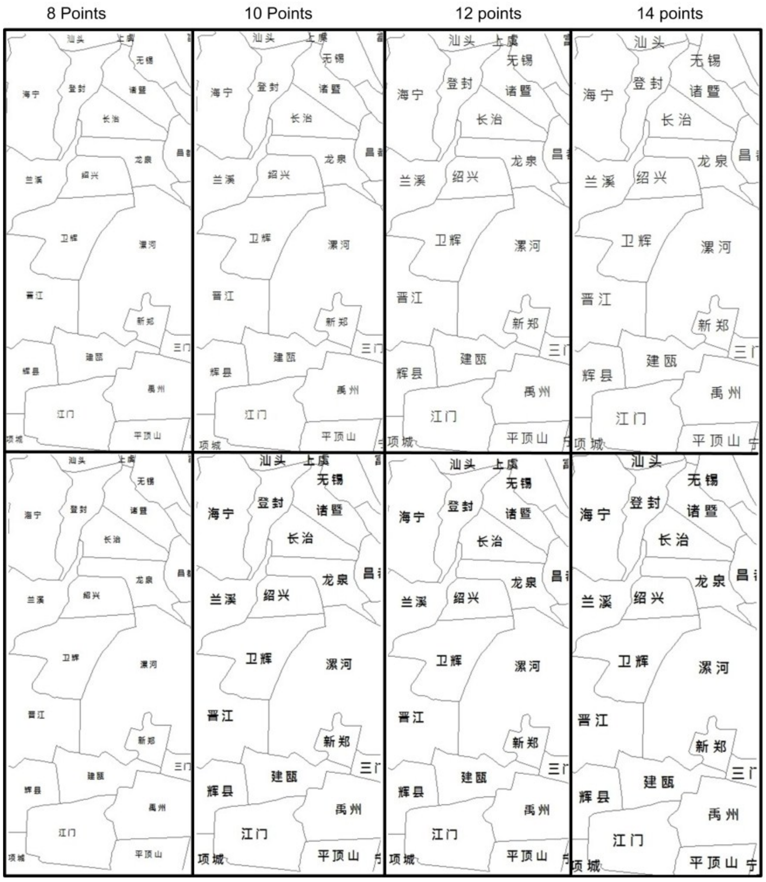


Figure 6.5: Size comparison of Chinese typography; the first row is normal size, and the second row is bold

6.4.1.1 Within lettering system size analysis

The mean reaction times of the test results are visualized in Figure 6.6. Single-factor ANOVA (one-way ANOVA) showed that the participants' efficiency in the three lettering systems was significantly influenced by different label sizes (F_L=3.407, P_L=0.019; F_A=18.940, P_A=0.000; F_C=6.129, P_C=0.001). However, there is an essential difference between the lettering systems regarding which sizes resulted in the most and least efficient searches. For the Latin participants, the most efficient size was 12 (M=7.597 s) and the least efficient was size 10 (M=11.671 s), with a slight difference from size 8 (M=11.514 s). For the Arabic participants, size 12 (M=5.417 s) was also the most efficient, whereas size 14 was the least efficient (M=15.006 s). The Chinese users were most efficient with the smallest size, 8 (M=5.684 s), and least efficient with size 10 (M=9.803 s).

An additional weight was added to the previous normal sizes (8, 10, 12 and 14) by making them bold. This enlarged the typographic size for all lettering systems and influenced the users' efficiency differently. The mean reaction time measurements regarding bold sizes are shown in Figure 6.6. This figure shows a higher efficiency of size 10, 12 and 14 (for the three lettering systems) in comparison to size 8. Like the normal size, Latin and Arabic users' efficiency of the four bold sizes were significantly different (F_L=4.010, P_L=0.008; F_A=4.610, P_A=0.004). However, the Chinese users' efficiency of bold sizes were not significantly different (F=1.333, P=0.265). For both Latin and Arabic lettering systems, the most efficient bold size was size 12 bold (M_L=5.913 s and M_A=6.326 s) and the least efficient size was 8 bold (M_L=8.323 s and M_A=9.697 s).

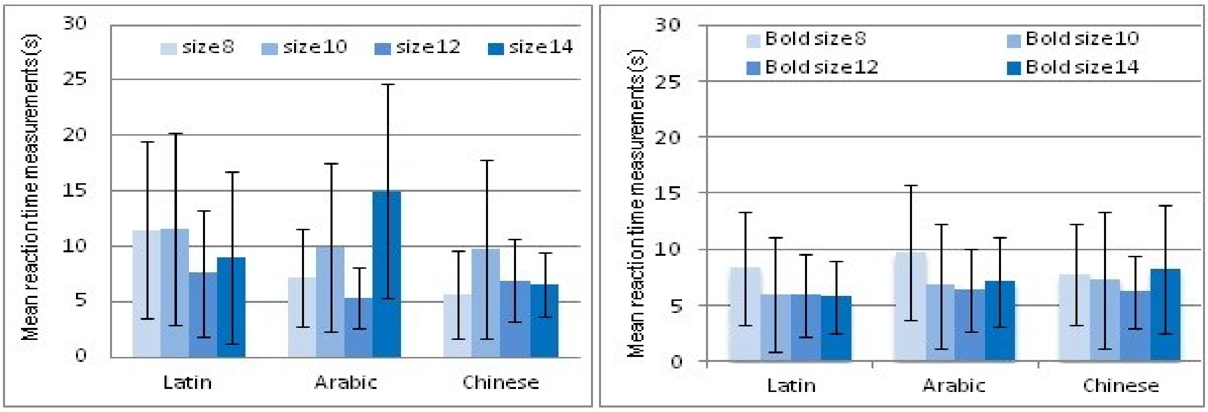


Figure 6.6: Mean reaction time measurements for the three lettering systems when different label sizes are used (left: normal labels; right: bold labels); the error bars represent the standard deviation

6.4.1.2 Between lettering systems size analysis

Using normal sizes showed different efficiency than using bold sizes for the same point measurement when comparing between the different lettering systems. First, Latin, Arabic and Chinese users' efficiencies were significantly different for Normal size 8 ($F_{8normal}$=13.951, $P_{8normal}$=0.000), with the highest efficiency for the Chinese lettering system (M_C =3.950 s) and the lowest for the Latin lettering system (M_L =7.936 s), but the efficiencies were not significantly different for bold size 8 (F_{8bold}= 1.727, P_{8bold}= 0.182). Second, both normal and bold size 10 were not significantly different ($F_{10normal}$= 0.782, $P_{10normal}$ = 0.459; F_{10bold}= 0.683, P_{10bold}= 0.507). Third, the users' efficiencies over the three lettering systems were significantly different for normal size 12 ($F_{12normal}$ =3.470, $P_{12normal}$ = 0.034), with the highest efficiency for Arabic (M_A=5.417 s) and the lowest efficiency for Latin (M_L=7.597 s). However, the efficiency for this size was not significantly different among the three lettering systems when presenting the labels in bold (F_{12bold}=0.153, P_{12bold}=0.858). Finally, the users' efficiencies across the three groups were significantly different for both normal and bold size 14 ($F_{14normal}$= 17.086, $P_{14normal}$= 0.000; F_{14bold}= 3.925, P_{14bold}=0.022). The highest efficiency of normal size 14 was for the Chinese lettering system (M_C=6.489 s), and the lowest efficiency of normal size 14 was for the Arabic lettering system (M_A=15.066 s). For bold size 14, the highest efficiency was for the Latin lettering system (M_L =5.746 s), whereas the least efficiency of bold size 14 was for the Chinese lettering system (M_C=8.243 s).

6.4.2 Shape

The application of the visual variable on labels is a critical issue, especially because the (different) lettering systems do not completely follow the rules of symbol design. Changing the shape of letters by applying a different font implies slight changes, as in Latin (serif versus sans serif, for example), or a massive change, as in Arabic (Diwani -curved lines- versus Koufi -angular lines). Furthermore, the fact that changing the shape at a fixed size will always be combined with a change in size, resulting in different consumptions of map space, must be considered for this study; two diverse fonts were applied for each lettering system using normal size 10 on a colourless background: a comparison was made between Arial and Times New Roman for Latin, between Tahoma and Times New Roman for Arabic, and between Microsoft JhengHei and SimSun for Chinese. The choices of the Arabic and Chinese fonts were based on their characteristic of resembling serif and sans serif fonts in Latin. However, these font families are not defined in both lettering systems.

In summary, shape A represents the sans serif font in Latin or its resemblances in the other lettering systems (Arial, Tahoma, Microsoft JhengHei), and shape B represents the serif font and its resemblances (Times New Roman and SimSun). Figure 6.7 shows an example of both shape groups over the three lettering systems.

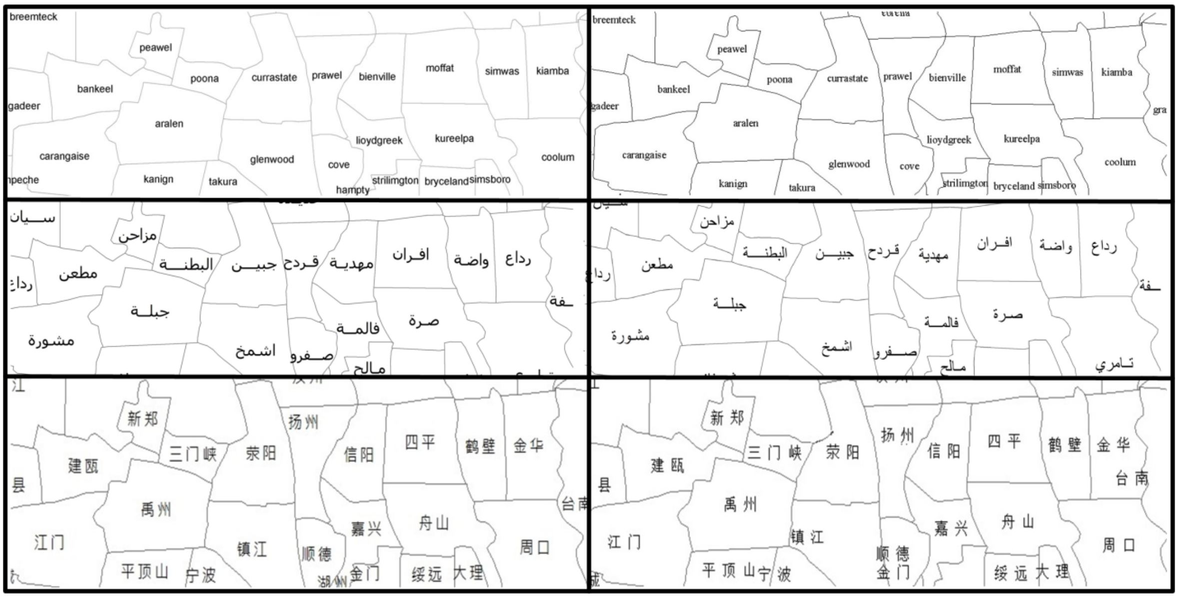


Figure 6.7: An example of both shape A (left) and shape B (right) for the three lettering systems

6.4.2.1 Within lettering system shape analysis

To have a clear view of the difference between shape A and shape B, a one-sample T test was used, which showed that the users' efficiencies were only significantly different with the Latin lettering system (F_L=12.068, P_L=0.001), where shape B (serif) was more efficient (M=6.786 s) than shape A (sans serif) (M=11.671 s). No significant difference was found with Arabic users (F=0.445, P=0.506) or with Chinese users (F=2.273, P=0.135). Figure 6.8 represents the mean reaction times for both shape groups over the three lettering systems.

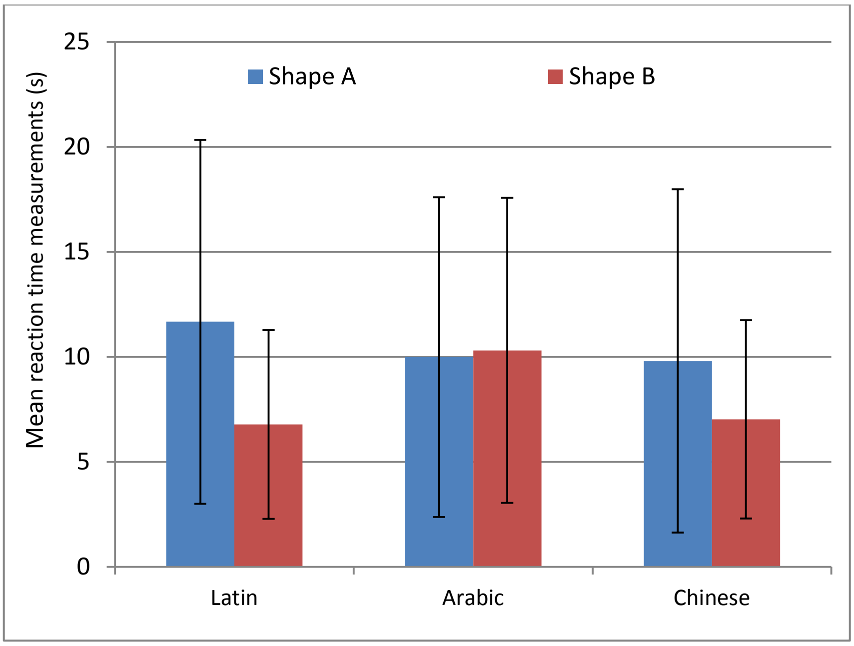


Figure 6.8: Mean reaction time measurements for users of different lettering systems and for different shapes; the error bars represent their standard deviation.

6.4.2.2 Between lettering systems shape analysis

The participants' efficiencies across the three lettering systems regarding shape A and shape B were tested by single-factor ANOVA, where shape A and shape B in the three lettering systems were tested. Shape A was not significantly different (F=0.782, P=0.459). Meanwhile, the participants' efficiencies with shape B (serif and its resemblances) were significantly different (F=5.964, P=0.003), with the highest efficiency for the Latin lettering system (M_L=6.786 s) and the lowest efficiency recorded for the Arabic lettering system (M_A=10.309 s).

6.4.3 Orientation

6.4.3.1 Overall label orientation

Over the thematic maps, the overall label orientation was studied in three categories: horizontal, tilted along the axis of the polygons' diagonal and mixed (horizontal if the label fits within the polygons border, tilted if not); this is illustrated in Figure 6.9. These orientations could be functional in some cases and could carry an aesthetic value of map design in other cases. A normal size 12 was applied to the three orientations and for the three lettering systems. The participants' mean reaction times regarding these different orientations are visualized in Figure 6.10.

6.4.3.1.1 Within lettering system overall orientation analysis

Single-factor ANOVA was used to identify the differences or similarities between the three orientation designs. The three orientations' efficiencies were significantly different with Latin users (F=7.924, P=0.001), Arabic users (F=13.830, P=0.000) and Chinese users (F= 4.380, P=0.014). The horizontal orientation of Latin and Chinese was the most efficient over the other two label orientations (tilted and mixed) in all lettering systems, but it was not the most efficient over the Arabic lettering system, as mixed orientation showed a higher efficiency.

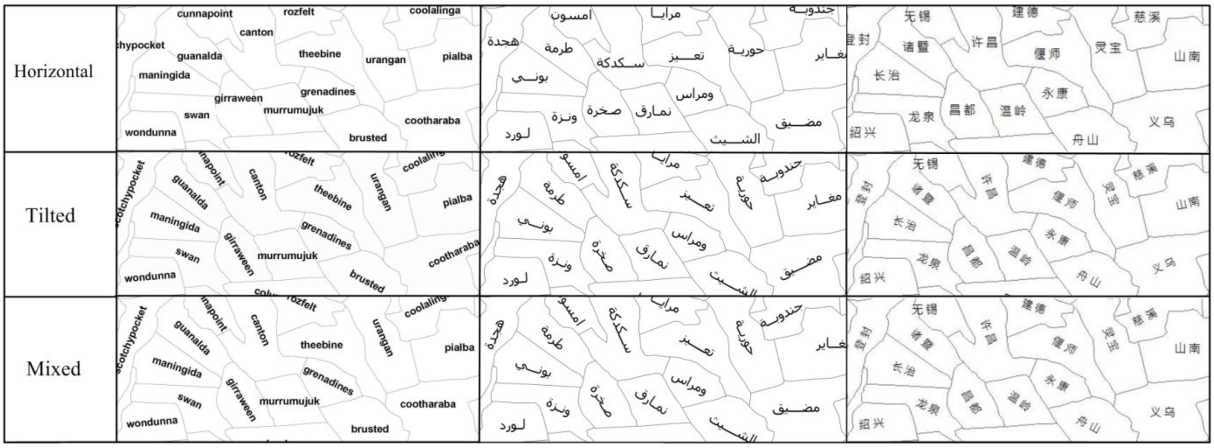


Figure 6.9: Small sketches of maps representing the overall label orientation

6.4.3.1.2 Between lettering systems overall orientation analysis

To study the orientation label design's influence on users' efficiencies between the three lettering systems, single-factor ANOVA was used to identify the differences or similarities between the three orientation designs. For the Latin, Arabic and Chinese lettering systems, the participants' efficiencies for the overall orientation design showed significant differences for horizontal, tilted and mixed orientation ($F_{horizontal}$=6.560, $P_{horizontal}$= 0.002; F_{tilted}= 14.147, P_{tilted}= 0.000; F_{mixed}= 4.104, P_{mixed}= 0.019). The highest efficiency for the horizontal orientation was for the Latin lettering system (M_L =9.043 s); for the tilted orientation, the highest user efficiency was for the Chinese lettering system (M_C =14.810 s); and for the mixed orientation, the highest user efficiency was for the Arabic lettering system (M_A=7.756 s).

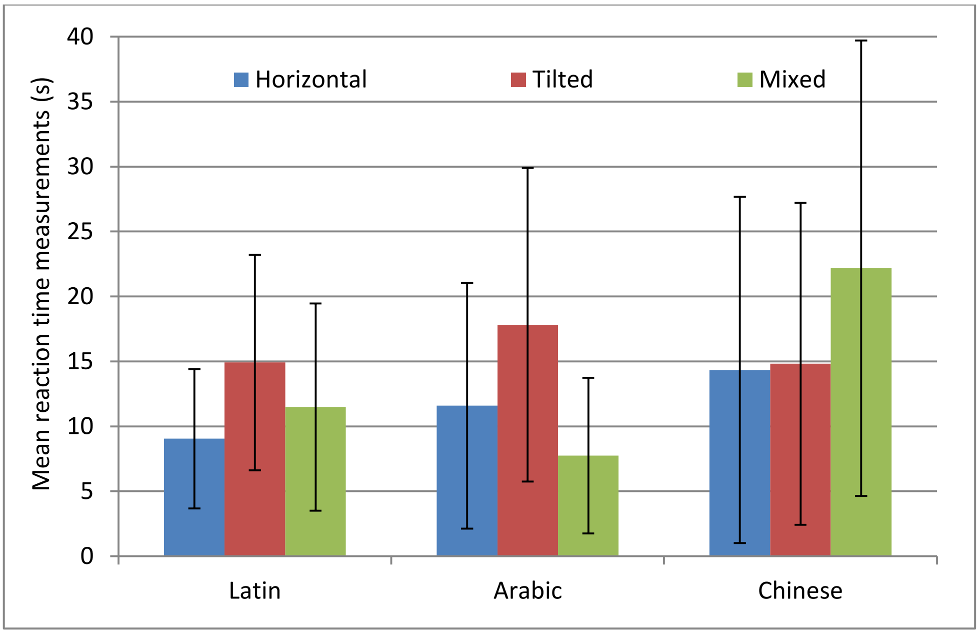


Figure 6.10: Mean reaction time measurements for users of different lettering systems and for different overall orientations; the error bars represent the standard deviation

6.4.3.2 Characters' orientation: italic

Although the typographic characters follow the same orientation within the word, we can still distinguish between two orientations: straight and italic (Deeb *et al.*, 2013). Therefore, making all characters italic with different variations in the labels' design was also implemented in the test. Four consecutive sizes (8, 10, 12 and 14) were combined with the italic orientation. In spite of the fact that italic typesetting is not commonly used in Arabic and Chinese text, it was implemented in the test for two reasons: first, the rules of cartographic text design do not completely match plain text design rules; second, map makers have the facilities to implement italic design in their maps for both functional and aesthetic reasons.

6.4.3.2.1 Within lettering system characters' orientation analysis

The mean reaction times for the four italic sizes are represented in Figure 6.11. Single-factor ANOVA was used to determine the similarities and differences of the four italic sizes within each lettering system. With the Latin lettering system, users' efficiencies were significantly different across the four italic sizes (F=5.063, P=0.002), and a similar result was obtained for the Chinese efficiencies (F=13.962, P=0.000). However, the Arabic lettering system did not indicate any significant difference for the use of italics with different sizes (F=1.115, P=0.344).

6.4.3.2.2 Between lettering systems characters' orientation analysis

When comparing between the different lettering systems, using italic text resulted in different efficiencies relative to using straight (normal) text for the same point measurement (see Figure 6.6 and Figure 6.11); the comparison was performed via single-factor ANOVA. First, users' efficiencies were not significantly different for italics among the different lettering systems with size 8 ($F_{8italic}$=0.260, $P_{8italic}$=0.771), whereas it was highly significantly different for straight (normal) size 8 ($F_{8normal}$=13.951, $P_{8normal}$= 0.000), see Figure 6.3, left. On the contrary, the use of italic size 10 for the three lettering systems was significantly different ($F_{10italic}$=9.434, $P_{10italic}$=0.000), with the highest efficiency for Latin (M_L=5.363 s) and the lowest efficiency for Arabic (M_A=9.313 s), whereas it was not significantly different when using straight (normal) size 10 ($F_{10normal}$= 0.782, $P_{10\ normal}$= 0.459). For the three lettering systems, the use of size 12 was significantly different for both italic ($F_{12italic}$=38.634, $P_{12italic}$=0.000) and straight (normal) ($F_{12normal}$=3.470, $P_{12normal}$= 0.034). The italic size 12 text has the highest efficiency for the Chinese lettering system (M_C= 3.045 s) and the lowest efficiency for the Arabic lettering system (M_A=8.586 s), whereas the straight (normal) size 12 text has the highest efficiency for Arabic (M_A=5.417 s) and the lowest efficiency for Latin (M_L=7.597 s). Finally, the users' efficiencies across the three groups were nearly significant for italic size 14 text ($F_{14italic}$=2.870, $P_{14italic}$=0.060) and highly significant when using straight size 14 text ($F_{14normal}$=17.086, $P_{14normal}$=0.000). As mentioned earlier, the highest efficiency for normal size 14 was for the Chinese lettering system (M_C=6.489 s), and the lowest efficiency for normal size 14 was for the Arabic lettering system (M_A=15.066 s).

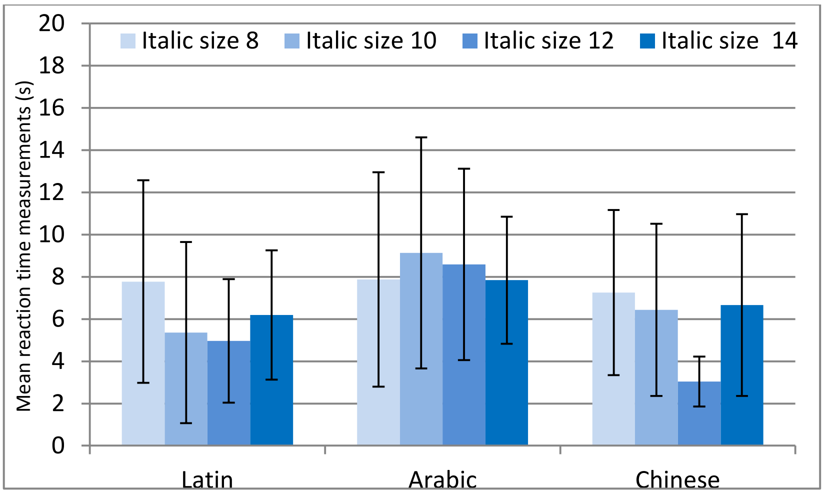


Figure 6.11: Mean reaction time measurements for italic orientation for users of different lettering systems over four consecutive sizes; the error bars represent their standard deviation

6.4.4 Texture

The definition of label texture is a critical issue because texture can be considered a semiotic system (Caivano, 1990). This study tested two elements of texture variation: the size (and the associated spacing) of the pattern elements (MacEachren, 1995) and the shape of the pattern elements. The distribution of labels across the map image, in combination with different label designs, creates a texture over the map face, where units' sizes and areas directly influence texture variations (Bertin, 1967). Figure 6.12 shows an example of the variation of two unit sizes (normal size versus bold size) in forming texture. In light of this texture variability, the hierarchical representation of label was studied using a combination of three consecutive sizes (8, 10 and 12) and two different fonts, resembling the serif (shape A) and sans serif (shape B) font families, in the three lettering systems (as explained above). The normal font design in the hierarchy was compared to the bold designs of both shape A and shape B. The three levels of hierarchy were represented on the same topographic map, where users had to search for a label in each level of hierarchy individually; the first level corresponded to size 14, the second level corresponded to size 12, and finally, the third level corresponded to size 10.

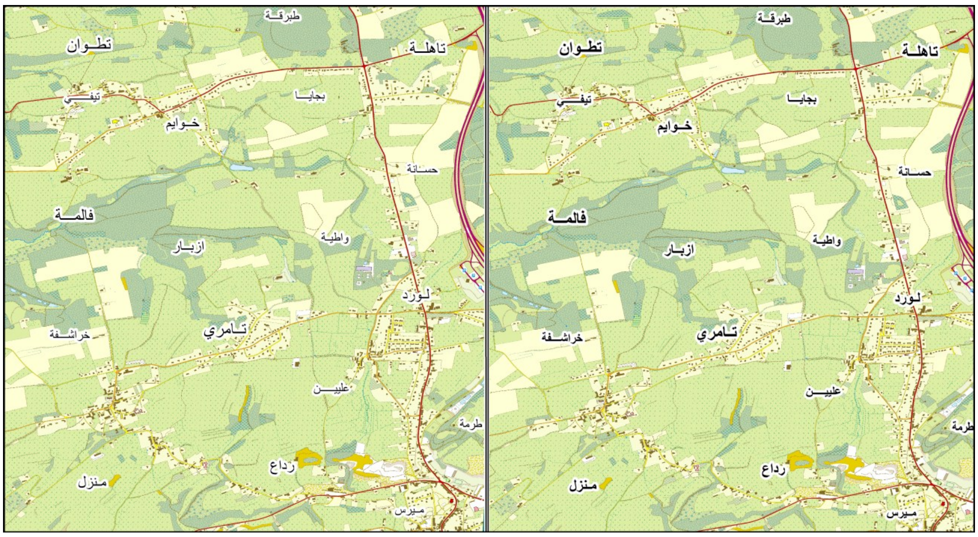


Figure 6.12: An example of two different textures: left, serif normal size hierarchy; right, bold size hierarchy (normal versus bold design, shape B)

6.4.4.1 Within lettering system texture analysis

A one-sample t test was used to compare the efficiency of shape A and shape B for normal sizes. Latin users' efficiencies were not significantly different for the first, second and third levels of hierarchy (F_1=0.008, P_1=0.930; F_2=0.264, P_2=0.608; F_3=0.484, P_3=0.488). Arabic users' efficiencies were also not significantly different for the first, second and third levels of hierarchy, (F_1=1.407, P_1=0.239; F_2=0.000, P_2=1.000; F_3=0.009, P_3=0.924). However, Chinese users' efficiencies for the first level of hierarchy were significantly different from those for the normal size (F=6.555, P=0.012), as shape A had a higher efficiency (M_C=5.443 s) than shape B (M_C=7.560 s). Conversely, users' efficiency was not significantly different for the second (F=0.000, P=1.000) and third levels of hierarchy (F=0.166, P=0.684).

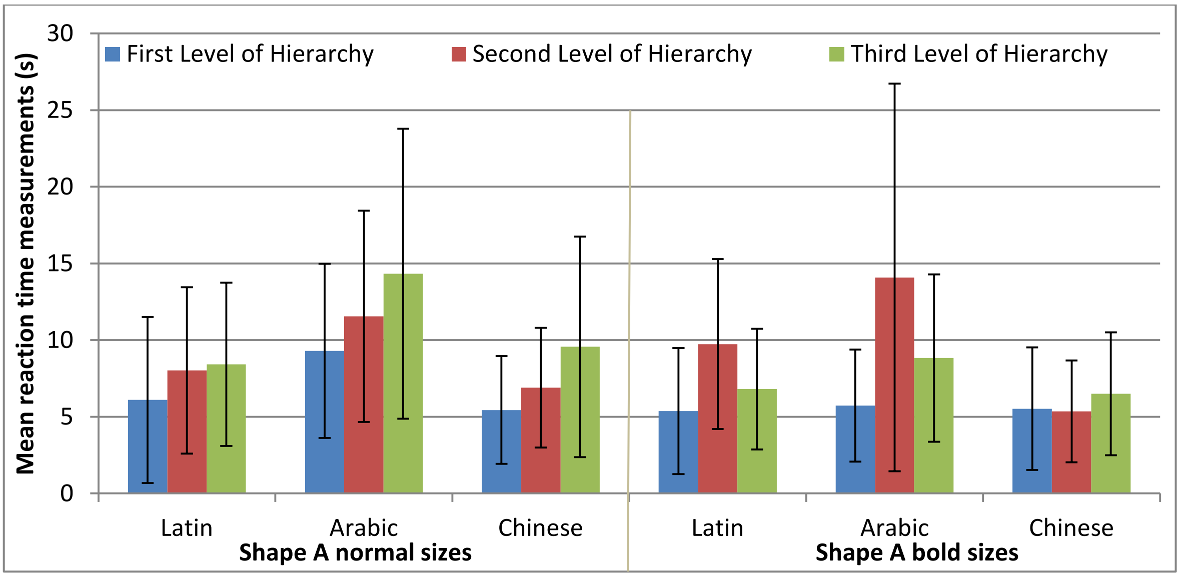


Figure 6.13: Mean values of users' reaction time measurements for the hierarchy of the three lettering systems for both the normal and bold sans serif-shape B-sizes, (left) normal and (right) bold; the error bars represent their standard deviation

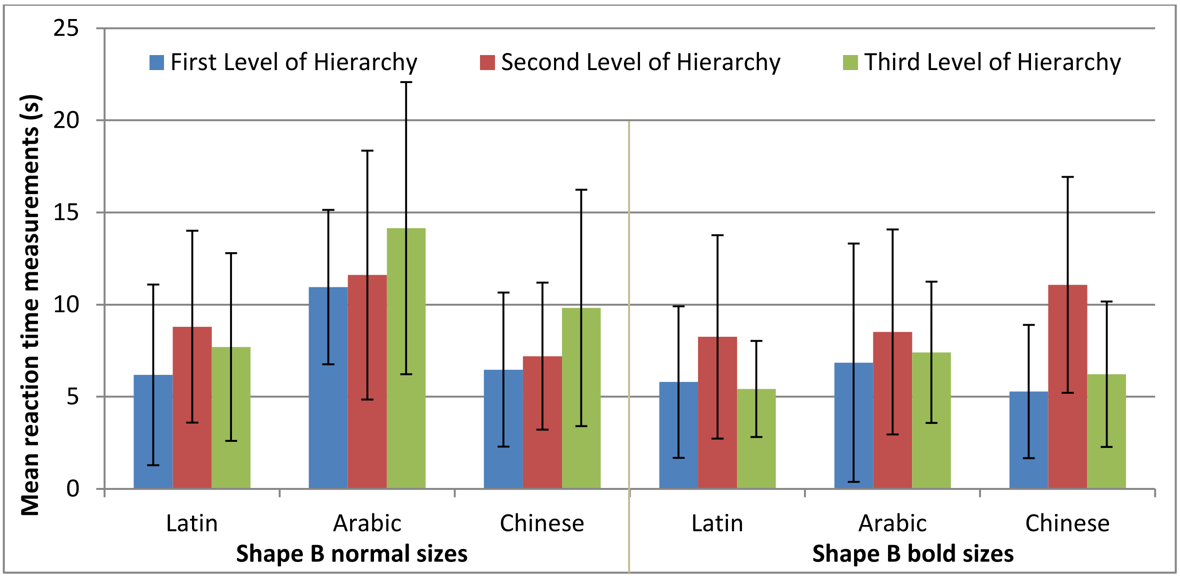


Figure 6.14: Mean values of users' reaction time measurements for the hierarchy of the three lettering systems for both the normal and bold serif-shape B-sizes, (left) normal and (right) bold; the error bars represent their standard deviation

Regarding the efficiency of shape A and shape B in the bold sizes, Latin users' efficiencies were not significantly different for the first and second levels of hierarchy (F_1=0.259, P_1=0.612; F_2=1.812, P_2=0.181), but they were significant for the third level of hierarchy (F_3= 4.310, P_3= 0.040), as shape A has a higher efficiency (M_L=5.418 s) than shape B (M_L=6.807 s). Arabic users' efficiencies were not significantly different for the first and third levels of hierarchy (F_1=1.691, P_1=0.197; F_3=1.987, P_3=0.162); however, they were significantly different for the second level of hierarchy (F2=6.274, P_2=0.014), as shape B had a higher efficiency (M_A=8.518 s) than shape A (M_A=14.083 s). Chinese users' efficiencies were not significantly different for both the first and third levels of hierarchy (F_1=0.442, P_1=0.502; F_3=0.509, P_3=0.477), but they were significantly different for the second level of hierarchy (F_2=26.646, P_2=0.000), as shape A had a higher efficiency (M_C=5.354 s) than shape B (M_C=11.074 s); see both Figure 6.13 and Figure 6.14.

6.4.4.2 Between lettering systems texture analysis

The participants' efficiencies for the three hierarchy levels were studied individually for each shape among the three lettering systems via single-factor ANOVA. The mean reaction times of each lettering system are presented in Figure 6.13 and Figure 6.14. Regarding shape A and normal size, the participants' efficiencies were significantly different among the three lettering systems for the first, second and third levels of hierarchy (F_1=11.464, P_1=0.000; F_2=7.499, P_2=0.001; F_3=8.206, P_3=0.000). The Arabic designs of hierarchy were the least efficient when using normal sizes with shape A over the three levels of hierarchy: (M_A=9.302 s), (M_A=5.681 s) and (M_A=14.325 s). The highest efficiencies for the first and second levels of hierarchy were for the Chinese lettering system: (M_C = 5.443 s) and (M_C =3.523 s), respectively, but for the third level of hierarchy, the highest efficiency was for the Latin lettering system (M_L =8.422 s). Additionally, regarding bold shape A, users' efficiencies for the first and third levels of hierarchy were not significantly different (F_1=0.026, P_1= 0.974; F_3= 2.894, P_3= 0.059). However, only for the second level of hierarchy were users' efficiencies among the three lettering systems significantly different (F_2=10.751, P_2= 0.000), as the highest efficiency was for the Chinese lettering system (M_C=7.201 s) and the lowest efficiency was for the Arabic lettering system (M_A=11.555 s).

When using Shape B for the three lettering systems, the normal size hierarchy showed significant differences between the first, second and third levels of hierarchy (F_1=7.329, P_1=0.001; F_1=5.919, P_1=0.003; F_1=9.414, P_1=0.000). For the first level of hierarchy, the highest efficiency was recorded for the Latin lettering system (M_L=5.367 s), and the lowest efficiency was recorded for the Chinese lettering system (M_C=7.56 s). The highest efficiencies for both the second and the third levels of hierarchy were recorded for the Chinese lettering system (M_C=5.354 s and M_C=6.493 s), and the lowest efficiencies for both the second and third levels of hierarchy were recorded for the Arabic lettering system: (M_A=14.083 s and M_A =8.828 s).

Similarly, Shape B with bold size significantly influenced users' efficiencies among the different lettering systems for the second level of hierarchy (F_2=4.273, P_2=0.016), with the highest efficiency for the Latin lettering system (M_L =8.249 s) and the lowest efficiency for the Chinese lettering system (M_C=11.047 s), but it did not influence participants regarding the first and the third levels of hierarchy (F_1=0.906, P_1=0.407; F_3=2.504, P_3=0.086).

6.5 Discussion

6.5.1 Label design efficiency

When applying visual variables onto the cartographic text, a user's efficiency is influenced differently according to each variable. The use of normal and bold fonts had an influence on the users' efficiencies over the three lettering systems. Introducing a variation in the labels' size influenced the users' efficiencies significantly; for the Latin, Arabic and Chinese users, size 12 was the most efficient for both the Latin and Arabic lettering systems, and size 8 was the most efficient for the Chinese lettering system. However, different bold sizes influenced users' efficiencies significantly for both Latin and Arabic users, as bold size 14 was the most efficient for the Latin lettering system, and bold size 12 was the most efficient for the Arabic lettering system. Furthermore, users' efficiencies for different bold sizes were not significantly different in the case of the Chinese lettering system. The experimental conditions showed that the Latin and Arabic lettering systems are more legible when using bold text, which made the efficiency of size use significantly different. Unlike the Chinese lettering system, the use of bold text eliminates the variations between the legibility of different sizes. This is highly probable because bold text decreases the point size variations in Chinese symbols as a whole, in comparison to Latin and Arabic text. Boldness played a basic role in Chinese letter system efficiency, as it makes it difficult to see the details of pictographic characters in certain sizes, with some of the Chinese characters completely unreadable in some font sizes. These results are in agreement with those of Huang *et al.* (2009) and Chan and Lee (2005), who found that different character sizes imply different degrees of legibility of plain text reading, whether it was on a computer display or a mobile display.

The comparisons between Shape A (sans serif font and its resemblances) and shape B (serif font and its resemblances) among the three lettering systems showed a significant difference only in the Latin lettering system. This issue might be explained by the fact that the serif and sans serif characteristics are only incorporated into Latin typography, even though typographic resemblance can be found in other lettering systems. The results support the work of Feldmann and Kreiter (2006), who found no significant difference for the sans serif font. However, further research needs to be undertaken to explore more shapes in the Arabic and Chinese lettering systems.

Under the experimental conditions, the overall orientation, including horizontal labels, tilted labels and mixed labels, significantly influenced the Latin, Arabic and Chinese users' efficiencies. This can be explained by the fact that reading horizontal labels is more efficient than reading labels with the other two orientation designs because it resembles plain text reading. However, for size 8, 10, 12 and 14, the orientation of the individual characteristics (italic) was significantly different for the Latin and Chinese lettering systems (italic size 12 was the most efficient for both lettering systems), but it was not significantly different for the Arabic lettering system. This can be explained by the fact that the italic orientation in Arabic typography is not commonly used. Therefore, the efficiency regarding the italic Arabic font was not found to be significantly different among the four sizes. Furthermore, the users' efficiencies for this typography were low in comparison with the average reaction times for Latin and Chinese text.

The users' efficiencies regarding the textural design varied between normal and bold fonts. In addition, there was no trend that could be attributed to the three levels of hierarchy. This could be linked directly to the characteristics of each lettering system. However, it is worth mentioning that the first level of hierarchy was not significantly different for both normal and bold shapes among all lettering systems. The research results do not agree with Bartz's (1970a) suggestion of using the typographic literature for map labelling. Our results also contradict those of Arditi and Cho (2005), who found that serif fonts are slightly more legible than sans serif fonts. However, their study focused on plain text, not on cartographic texts. Therefore, the results suggest the importance of considering the cartographic text functions as Fairbairn (1993) defined, in addition to considering the lettering system's nature and semiotics.

6.5.2 Cross cultural comparison

Latin, Arabic and Chinese lettering systems are inherently linked with their own design variations. However, when applying some of the visual variables on cartographic labels in these systems, the users' efficiencies showed no significant difference, whereas other variables influenced the users' efficiencies to a significant degree. This variation can be explained not only by the nature of the lettering systems themselves but also by the changes that each variable presented over the characters. In comparison with the normal and bold sizes, the results showed that the normal sizes influence the efficiency of the three lettering systems for normal size 8, 12 and 14, where the most efficient 8 normal size was for the Chinese lettering system, the most efficient 12 normal size was for the Arabic lettering system and the most efficient 14 normal size was for the Chinese lettering system. In addition to the normal sizes, bold sizes influenced users' efficiencies among the three lettering systems for bold size 14, with the highest efficiency for the Latin lettering system. It is worth mentioning that both normal and bold size 10 were in excellent agreement, having no influence on the three lettering systems, in contrast with normal and bold size 14, which had significant differences among the three lettering systems. In addition to size influence, shape also influenced users' efficiencies differently among the three lettering systems when comparing two different designs of shapes: sans serif and its resemblance and serif and its resemblance. The first group (Shape A) did not show any significant difference among the three lettering systems, whereas the second group (Shape B) showed a significant difference, with the highest efficiency for the Latin lettering system and the lowest efficiency for the Arabic lettering system. However, we do not yet have a direct experimental explanation to clarify the details of this issue under the experimental and stimulus conditions. Further research needs to be undertaken in order to explain the shape's influence regarding the three lettering system, as there are different standards to describe the shapes in each lettering system.

Considering the overall word orientation design (horizontal, tilted, mixed), the three lettering systems showed agreement. However, the users' efficiencies were significantly different among the three orientation designs. When considering the italic character orientation, the efficiency was not significantly different between the smallest and largest size (8 and 14), while it was significantly different when using size 10 and 12, as Latin had the highest efficiency for italic size 10 and Chinese had the highest efficiency for italic size 12. Here, there is a direct link between the italic orientation efficiency and the characters' sizes. To explain these results, we suspect that italic text can hide some of the characters' details and would make words practically unreadable in Arabic. Therefore, it is crucial to choose the right size when using italic text.

The three lettering systems influenced users' efficiencies significantly at the first, second and third hierarchical levels when using normal sizes for both shape A and shape B. However, when using bold sizes, the only significant difference was recorded for the second level of hierarchy for both shape A and shape B, where the highest efficiency was for the Chinese lettering system when using shape A and for the Latin lettering system when using shape B. The Arabic lettering system, among the three levels of hierarchy was the least efficient when using normal sizes for shape A and the least efficient for the second and third levels of hierarchy when using shape B. It is not clear at this level of analysis and with these experimental conditions whether this significant difference occurs because of the lettering systems themselves or because of the definition that we set for shape A and shape B. Because texture is formed in combination with shape, obtaining more information about the effects of shape on users' efficiencies in the future will allow us to refine our prediction in due course.

6.5.3 Limitation and Generalizability

The results are consistent with our definition of size, but they are limited to both the used point size and the screen size. Therefore, our results are limited to screen maps that have a similar map scale or screen maps that preserve the ratio between labels' sizes and map scale. Additionally, these results can be generalized to combine interactive maps that only use pan operation, which also preserves the ratio between map scale and display medium.

In light of our classification of typographic shape, our findings can be generalized to cover the type groups that resample what we have called shape A and shape B. Group A represents the type design whose additional lines are the least possible, and shape B is the group that contains many additional lines. This generalizability is consistent with the results of Feldmannand and Kreiter (2006), who tested three sans serif fonts (Arial, Univers and Frutiger) and did not find a significant effect on users' efficiencies among the three fonts. However, some shapes cannot be classified in either of the mentioned groups because they have outstanding characters, such as Koufi in Arabic, which is similar to the oriental geometrical ornament, and Forte in Latin, which is similar to handwriting. Additionally, considering the orientation of the labels, the experimental conditions confine the application of the overall label orientation to its relation to the district topology. Label orientation was not studied in regard to point and linear features.

Furthermore, cartographic text texture was only studied by combining both shape and size. Other categories of texture can be formed by integrating orientation into the design.

This study does not cover all map display sizes, which can play a significant role in users' efficiencies towards map labelling in relation to different media sizes. Moreover, none of the lettering systems were combined with their diacritic marks; as a result, it is highly expected that users' efficiencies will be affected by the typographic design that includes diacritic marks, especially in languages that use them very often (such as Czech).

6.6 Conclusion

This research attempted to address solutions for some of the cartographic text design issues (and, more specifically, the labels) regarding both user efficiency and cultural influences. Two major points were examined. First, user efficiency differs when applying different visual variables on the cartographic text design. Second, user efficiency is influenced by lettering system designs. The first point is highly important when making monolingual maps that include the Latin, Arabic or Chinese lettering systems. The second point concerns the multilingual cartographic text design and indicates the average efficiency for each visual variable. These results can be used as a guide for map designers. They also allow designers to equalize the importance of multilingual labels over the map by applying the variables that have no significant influence over the three lettering systems. Conversely, they may be used to categorize the importance of multilingual labels by using different variability, which provide higher or lower user efficiency for each lettering system; thus, it is highly probable that users will read the label with the highest efficiency first, and the other(s) will be categorized as supported information. Further experiments need to be carried out to rectify our suggestion.

Because the cartographic products can be presented on different display media, it would be worthwhile to obtain more experimental information about the cross-cultural cartographic text design by considering interactive maps with different display devices.

6.7 Appendix-1

Map	Map type	Typographic variable	Latin target	Arabic target	Chinese target
1	Thematic map	Size 10, shape sans serif	Aralen	جبلة	禹州
2	Topographic map(rural area)	Third level of hierarchy , bold sans serif	Strueth	بجايا	项城
3	Topographic map(urban area)	First level of hierarchy, bold sans serif	Aroona	انزيغ	商丘
4	Thematic map	Orientation, mixed	Peawel	مزاحن	新郑
5	Topographic map(urban area)	Second level of hierarchy, sans serif	Morant	خوايم	格尔木
6	Thematic map.	Size 8	Moffat	افران	四平
7	Topographic map(urban area)	First level of hierarchy, bold serif	Marietta	تامري	焦作
8	Thematic map.	Size 10, shape serif	Oclame	مناخة	桐乡
9	Topographic map(rural area)	First level of hierarchy, serif	Danpire	صفرو	嘉义
10	Topographic map(urban area)	Second level of hierarchy, bold sans serif	Marshall	مكناسة	乐清
11	Thematic map.	Size 14 italic	Takura	اموجور	宁波
12	Topographic map(rural area)	Second level of hierarchy, serif	Greenfield	مطعن	瑞安
13	Thematic map.	Size 8 italic	Coolum	تامري	周口
14	Topographic map(rural area)	First level of hierarchy, bold sans serif	Aroona	انزيغ	商丘
15	Thematic map.	Size 12 bold	Agadeer	سيان	晋江
16	Topographic map(rural area)	First level of hierarchy, bold serif	Marietta	تامري	焦作
17	Thematic map.	Size 10 bold	Millner	ججاز	漯河
18	Topographic map(urban area)	Second level of hierarchy, bold serif	Gadseden	منزل	漯河

19	Thematic map.	Size 14	Simwas	واضة	鹤 壁
20	Topographic map(urban area)	First level of hierarchy, serif	Danpire	صفرو	嘉 义
21	Topographic map(urban area)	Third level of hierarchy, san serif	Peawell	وزان	偃 师
22	Thematic map.	Size 12 italic	Poona	البطنة	三 门 峡
23	Topographic map(urban area)	Third level of hierarchy, bold serif	Bunnel	رقاب	宁 波
24	Thematic map.	Size 10 italic	Cookely	نمارية	阿 里
25	Topographic map(rural area)	Second level of hierarchy, bold serif	Gadesten	منزل	漯 河
26	Topographic map(urban area)	Third level of hierarchy, serif	Fanidn	واطية	扬 州
27	Topographic map(rural area)	First level of hierarchy, sans serif	Deliman	قيروان	台 南
28	Thematic map.	Size 14 bold	Azmeer	مودية	舞 钢
29	Topographic map(urban area)	Third level of hierarchy, bold sans serif	Strueth	بجايا	项 城
30	Topographic map(rural area)	Second level of hierarchy, bold sans serif	Marshall	مكناسة	乐 清
31	Topographic map(rural area)	Third level of hierarchy, serif	Fandin	واطية	扬 州
32	Thematic map.	Orientation, horizontal	Bankeel	مطعن	建 瓯
33	Topographic map(rural area)	Third level of hierarchy, bold serif	Bankeel	رقاب	宁 波
34	Topographic map(rural area)	Second level of hierarchy, sans serif	Morant	خوايم	格 尔 木
35	Thematic map.	Size 8 bold	Bienville	مهدية	信 阳
36	Topographic map(rural area)	Third level of hierarchy, sans serif	Peawel	وزان	偃 师
37	Thematic map.	Orientation, tilted	Pialba	مغاير	山 南
38	Topographic map(urban area)	Second level of hierarchy, serif	Greenfield	مطعن	瑞 安
39	Thematic map.	Size 12	Gympie	سمسرة	漳 州
40	Topographic map(urban area)	First level of hierarchy, sans serif	Deliman	قيروان	台 南

6.8 References

Al-Harkan, I. M. & Ramadan, M. Z. 2005. Effects of Pixel shape and color, and matrix pixel density of Arabic Digital Typeface on characters' legibility, International Journal of Industrial Ergonomics, 35, 625-664.

Arditi, A. & Cho, J. 2007. Letter case and text legibility in normal and low vision, Vision Research, 47(19), pp.2499-2505.

Arditi, A. & Cho, J. 2005. Serif and font legibility, Vision Research, 45 (23), pp.2926-2933.

Azmi, A. & Alsaiari, A. 2010. Arabic Typography: a Survey, International Journal of Electronical & computer sciences, 9 (10), pp.16-22.

Bartz, B. S. 1970b. Experimental Use of The Search Task in The Analysis of Type Legibility in Cartography, The Cartographic Journal, 7(2), pp.103-113.

Bartz, B. S. 1970a. An Analysis of the Typographic Legibility Literature, Assessment of Its Applicability to Cartography, The Cartographic Journal, 7(1), pp.6-17.

Bernard, M. L., Chaparro, B. S., Mills, M. M. & Halcomb, C. G. 2003. Comparing the effect of text size and format on the readability of computer-displayed Times New Roman and Arial text. International Journal of Human-Computer Studies, 59(16), pp.823-835.

Bertin, J. 1967. Semiology of Graphics: Diagram, Networks, Maps, Translated by William J. Berg 2011, Esri press,California.

Caivano, J. L. 1990. Visual texture as a semiotic system, Semiotica, 80(3/4),pp.239-252.

Chan, A. H. S., & Lee, P. S. K. 2005 Effect of display factors on Chinese reading times, comprehension scores and preferences, Behaviour & Information Technology, 24(2), pp.81-91.

Deeb, R., Ooms, K., Brychtova, A., Van Etvelde, V. & De Maeyer, P. In press. Background and foreground interaction: complementary colours' influence on search task, Color Research and Applications. DOI: 10.1002/col.21920.

Deeb, R., Ooms, K., Vanopbroeke, V., & De Maeyer, P. 2013. Evaluating the Efficiency of Typographic Design: Gender and Expertise Variation, The Cartographic Journal, 51(1), pp.75-86.

Deeb, R., Ooms, K., and De Maeyer, P. 2012. Typography in the Eyes of Bertin, Gender and Expertise Variation, The Cartographic Journal, 49(2), pp.176-185.

Fairbairn, D. J. 1993. On The Nature of Cartographic text, The Cartographic Journal, 30(2), pp.104-111.

Feldmann H-U & Kreiter N. 2006. Neuaufbau der schweizerischen, Landeskarte: Inhalt und Kartengrafik. Kartographische Nachrichten, 3, pp.115–121.

Gerber, R. V. 1981.Competence and Performance in Cartographic Language, The Cartographic Journal, 18 (2), pp.104-111.

Hersch, R. D. 1994. Font rasterization: the state of the art. In From object modelling to advanced visual communication, pp.274-296. Berlin Heidelberg.

Huang, D-L., Rau, P-L. P., & Liu Y. 2009. Effects of font size, display resolution and task type on reading Chinese fonts from mobile devices. International Journal of Industrial Ergonomics, 39(1), pp.81-89.

Kraak, M. J. & Ormeling, F. 2010. Cartography, Visualization of Spatial Data, 3rd edition, Harlow: Person Prentic Hall.

MacEachren, A. M. 1995. How Maps Work, Representation, Visualization and Design, the Guilford press, NY.

Phillips, R. J. 1981. Search for Target in a Random Arrangement of Names: an Eye Fixation Analysis, Canadian Journal of Psychology, 35(4), pp.330-346.

Phillips, R. J., Noyes, E. & Audley, R. J. 1978. Searching for names on maps. The The Cartographic Journal, 15(2), pp.72-77.

Phillips, R. J., Noyes, L.& Audley, R. J. 1977.The legibility of type on maps. Ergonomics, 20(6), pp.671-682.

Taylor, R. M. & Hopkin, V. D. 1975. Ergonomic principles and map design. Applied ergonomics, 6(4), pp.196-204.

Trenkle, J., Gillies, A., Erlandson, E., Schlosser, S., & Cavin, S. 2001. 'Advanced in Arabic Text Recognition'. In: the proceeding of symposium on Document Image Understanding Technology, pp.158-168, Columbia: Maryland.

van den Worm, J. 2001. 'Web Map Design in Practice', In: Kraak, M.J. and Brown, A. Web Cartography, Taylor & Francis, chapter 7, pp.87-107, London.

Woodward, D. 1987. 'The Manuscript, Engraved, and Typographic Traditions of Map Lettering'. In: Woodward, D. eds. Art and Cartography, The University of Chicago Press. Chapter 6, pp.174-212, Chicago.

7

Assessment of Map Users' Search Strategies Using Multi-Package Software for Scanpath Analysis

Modified from: Deeb, R., Ooms, K., Popelka, S., Van Etvelde, V. & De Maeyer, P. 2015b. Assessment of map users' search strategies using multi-packages software for scanpaths analysis. 27th International Cartographic Conference, Rio de Janeiro (accepted contribution).

Preface

The eye tracking data proved to be a trustworthy data to asses map usability issues. As both the eye tracking quantitative and qualitative data were used to analyse different applications. This chapter describes a user study in which users' had to locate a target label. Map users' search strategies were recorded by SMI Eye tracking device. The study combined multiple package open source software that can analyze the sequences of the users' scanpaths by comparing strings. The map was divided into several Areas Of Interest (AOI) to which a letter-code was assigned. The resulting strings (subsequent letter-codes of AOIs that were fixated) can be used to define the similarity of the scanpaths. The software packages OGAMA and EyePatterns were used to visualize and analyze the users' search patterns. Different characteristics of the users (expertise, gender) were taken into account during the analyses. Substantial qualitative data analyses were provided by the combination of these software' applications. Besides, a number of search patterns were presented. This work demonstrates the use of multi-packages software for scanpaths analysis.

7.1 Introduction

The last decades witnessed an increasing interest in the eye tracking technique, mainly from the computer-graphics community due to the rich set of visual productions, displays, stimuli, and applications that has been introduced. Eye tracking research has statistically tackled eye tracking metrics as gaze parameter, scanpath length, and other quantitative data (Duchowski, 2002; Yang, 2002; and Rayner, 1998). These quantitative data was effective for analysing and interpreting a certain applications. However, a few studies have considered the qualitative analysis of scanpaths and scanpath clustering (Andrienko *et al.*, 2012 Adriano *et al.*, 2009; Augustyniak and Mikrut, 2006; Ooms *et al.*, 2012).

Scanpaths analysis is necessary to clarify cognitive processes. Scanpaths are dependent on subject, stimulus, and task which drive the sequences (Brandt & Strak 1997; Strak & Ellis 1981). Treating scanpath statistically (Augustyniak, 2006; Augustyniak, 2003) does not reveal much cognitive information about the technique that users perform while they search for a target. Both bottom-up and top-down stimuli processing are necessary to explain the complexity of object cognition (labels in this study). Since we cannot determine whether perceivers first interpret the whole image or first interpret the parts (Matlin, 2003), It is crucial to know the details of the processes that direct the visual attention while searching for a target on the map. Brandt and Strak (1997) tried to determine the degree of similarity of saccades using string editing analysis in order to determine whether the eye movement patterns give insights to a cognitive model. They were able to model the users' scan paths in seven rectangular AOIs. Byrne *et al.* (1999) used eye tracking to test the visual search on a very basic schema of one column that had either one letter target, one number target or none target for distracters in the column cells. They concluded for their stimuli that search was primarily top to bottom and rarely appeared to be random. Similarly, Hornof and Halverson (2003) examined search path for targets over six parallel discrete columns that had multiple uppercase letters in the individual cell. They concluded eight different search strategies from their stimuli. Although the mentioned study used eye tracking to describe the visual search on different schemas, and they provided un insight about the verity of search strategies, but none of them can explain how users' search on a map, because the shapes and colours on the map are not systematic, unlike what they used for their stimuli.

Searching for targets on static maps is an important map designing issue. When map designer know the way in which users' search for targets on maps, they can take that into account to produce maps which content is easier to process and interpret. The earliest attempt that investigated search strategies on maps and its influencing factors was conducted by Phillips *et al.* (1978). They concluded that the search strategies have a random influence on the number of fixated names and thus the scanpath strings. However, reading a map does not follow the same patterns as reading a plain text (left to right, right to left, and top to bottom). In addition to that, there is no predetermined path for the eyes to follow and the gaps between labels are unexpected and rather larger than on a page of text (Phillips, 1981). Furthermore, Wolf (1994) tried to explain our ability to find a desired visual stimulus in a normal, continues visual scene. As he concluded that the attention is derived by either bottom-up activation or top-down activation. Moreover, Lleras and Mühlenen (2003) clarified that visual search strategies can vary between users. They discovered that different search approaches were followed. The search strategies were divided into either a systematic approach, as users geometrically scanned the map (top to bottom, bottom to top, right to left, and left to right), or an intuitive approach, as the users intuitively searched the map image. This makes searching for labels on a map far more unpredictable than the words on plain text. Besides that, personality characteristics shall be considered because it affects on the task behaviours is not clear (Gellatly, 1996). How the search strategies can be analyzed using the qualitative characteristics of scanpaths has not been investigated previously. The goal of this chapter is to analyze search scanpaths by using a combination of several software packages. The results show that these software packages are ideal tools to address the search patterns and the variability between different users' groups. An eye tracking experiment and its analysis are reported in the next sections.

7.2 Study design

7.2.1 Apparatus

A controlled experiment was conducted in the eye tracking laboratory of the Department of Geography, Ghent University, Belgium. The laboratory is equipped with and SMI RED 250 eye tacking device with Sampling rate of 120Hz (SensoMotoric Instruments GmbH, Germany). The stimuli were displayed on a 22-inch DELL (resolution: 1680×1050 pixels; refresh rate: 60 Hz).

7.2.2 Participants

A trail of 12 maps was presented to 31 subjects. Due to the fact that user characteristics are crucial to determine (Haklay & Nivala, 2010; Nielsen, 1993; Rubin & Chisnell, 2008; Duchowski, 2007), the study has a between user design. The subjects consisted of 14 females and 17 males considering users' gender on the one hand and on the other hand they consisted of 15 experts and 16 novices considering their level of expertise. The novices were bachelor students who had just started their education in Geography and Archaeology with an average age of 19 years. The experts had obtained at least a Master degree in Geography and Geomatics and they work with maps on a daily basis. Their average age was 26 years.

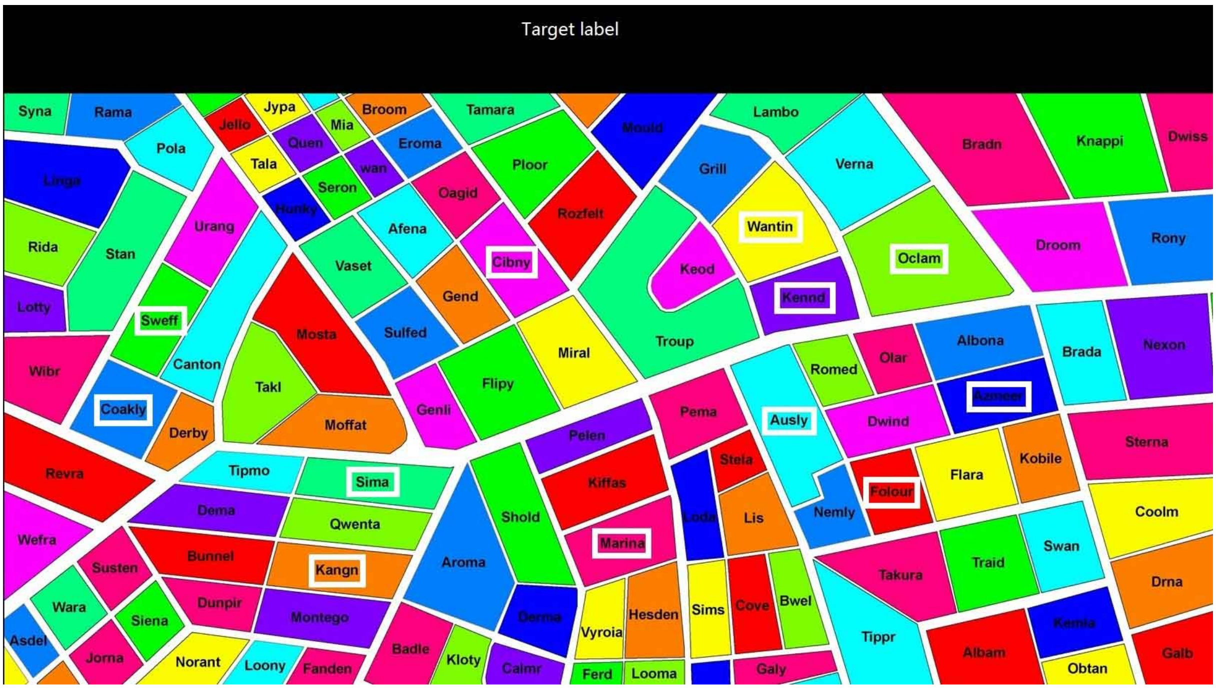


Figure 7.1: The targets distribution over the stimuli dimension. Targets are surrounded with white rectangles

7.2.3 Task and stimuli

A number of 12 target labels were associated with search tasks. For each stimulus, subjects were asked to find one target label which was distributed differently over the map face (see Figure 7.1). Black target labels were depicted over the map face, while white targets were asked at the top of map image. Each map has a target label placed on a different background colour compared to other maps. Therefore, 12 different colour hue were involved (0, 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330). The stimuli tried to cover the colours spectrum to mimic the colour relationships in the normal cases of map design (Cleveland & McGill, 1983). Subjects were asked to locate the target label in the map image by fixating on it. Since the target label was visualized on top of the actual map image, the search path is thus recorded starting from the target label on top of the map image and ending with the fixation on the actual label in the map image. When a subject finds the target label in the map, the next stimulus automatically appears.

7.2.4 Procedures

Data were recorded in the SMI eye tracking device. Then the row data was collected and exported from SMI BeGaze. Afterwards the row data was imported, processed, and partly analyzed in OGAMA (Open GAze and Mouse Analyser). Then, these processed data was imported to EyePatterns (West *et al.,* 2006), the sequence analysis tool for string analysis and pattern visualization. Search patterns were created in Eyepatters then they were redrawn for better visualization. Later, Search clusters were explained using OGAMA. See Figure 7.2.

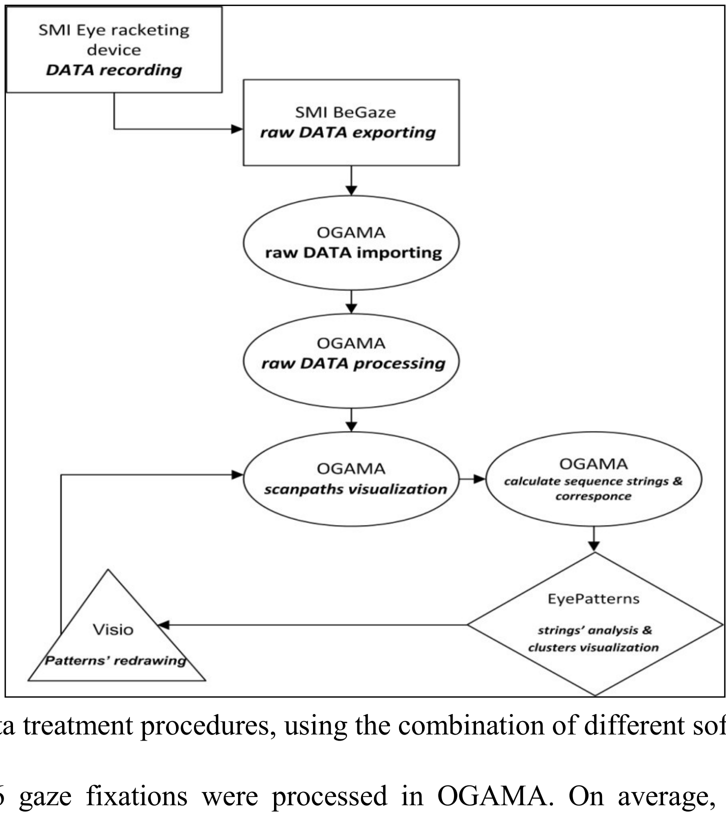


Figure 7.2: Data treatment procedures, using the combination of different software packages

In total, 33976 gaze fixations were processed in OGAMA. On average, 94.4 fixations were processed for each subject per stimulus, see Table 7.1. The variation between users' performance is listed in Table 7.1. Areas Of Interest were drown over the map face (Figure 7.3) and These fixations were transferred into trails of alphabetical letters, which correspond to their location on AOIs instead of their occurrence number. This process drew an alphabetical scanpath strings. EyePatterns treated the fixation sequence as expanded sequences automatically, so it considered all the subsequent letters as one letter. Consequently, successive fixations that occurred within the same AOI get only one letter-code assigned to them. Thus, only one saccade occurs between two AOIs. This makes tracing and visualising search systems simpler for each stimulus. For this reason AOIs were chosen carefully taking into account the overall shapes of the main features in the studied map, illustrated in Figure 7.3. A thorough explanation is provided in section 7.3.1.

Table 7.1: Users' fixation average per stimulus and fixation sum considering their characteristics

Experts	Gender	Average fixation	Fixation sum	Novices	Gender	Average fixation	Fixation sum
				Participant 16	Female	66.3	796
Participant 1	Female	189.3	2272	Participant 17	Female	83.9	1007
Participant 2	Female	54.9	659	Participant 18	Female	47.9	575
Participant 3	Female	131	1572	Participant 19	Female	101.3	1216
Participant 4	Female	75.3	904	Participant 20	Female	114.2	1371
Participant 5	Female	93.1	1117	Participant 21	Female	177.9	2135
Participant 6	Female	160.6	1928	Participant 22	Female	93.1	1117
Participant 7	Female	97.4	1169	Participant 23	Male	92.8	1114
Participant 8	Male	92.2	1106	Participant 24	Male	122.9	1475
Participant 9	Male	69.7	837	Participant 25	Male	82.5	990
Participant 10	Male	92.6	1111	Participant 26	Male	103.9	1247
Participant 11	Male	85	1020	Participant 27	Male	95	1140
Participant 12	Male	105.7	1269	Participant 28	Male	160.6	1927
Participant 13	Male	88.6	1064	Participant 29	Male	70.3	843
Participant 14	Male	96.5	1158	Participant 30	Male	62.3	748
Participant 15	Male	90.1	1082	Participant 31	Male	59.3	712

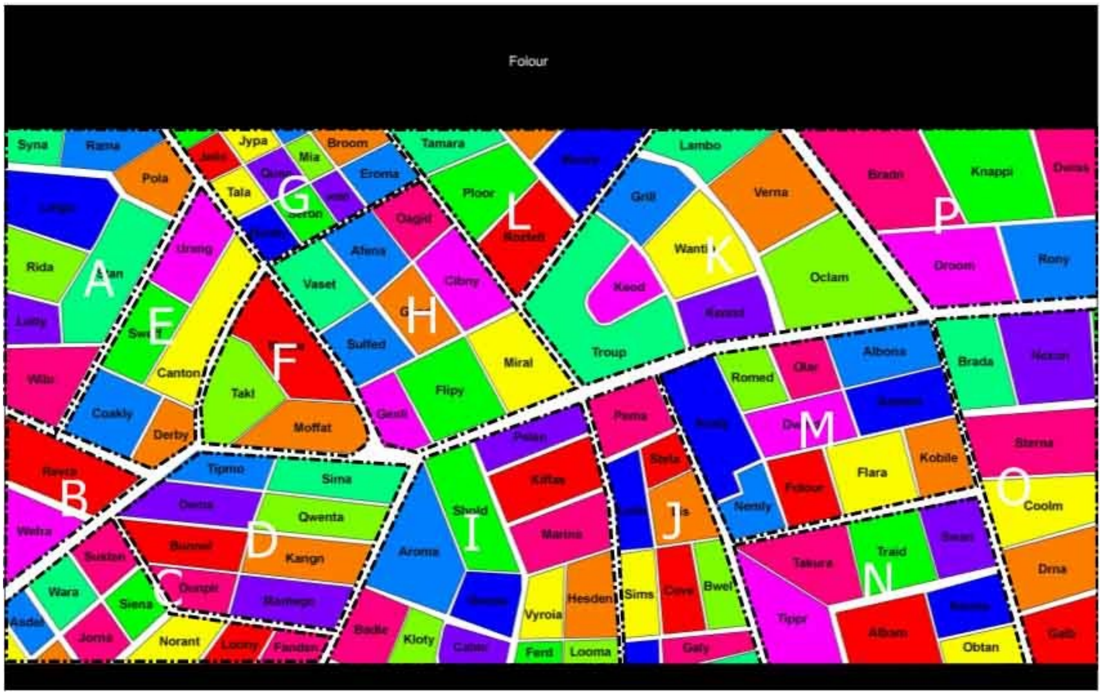


Figure 7.3: Areas of Interests' design according to the spatial structure of the stimuli

7.3 Results & Discussion

Examples of scanpaths from different participants are illustrated in Figure 7.4. The first two participants show two opposite search strategies, searching on the map with subject 4 (S4) starts from the upper left to the lower right and the other scanpath with S18 starts from the upper right towards the lower left. This could be described as a systematic search because the subject scans the map in defined pattern. But in other cases such as the examples with S9 and S13, the search system is more unstructured and cannot be described as a systematic search. Because of this we cannot describe the search path so easily; in this case users move their eyes in all direction arbitrary.

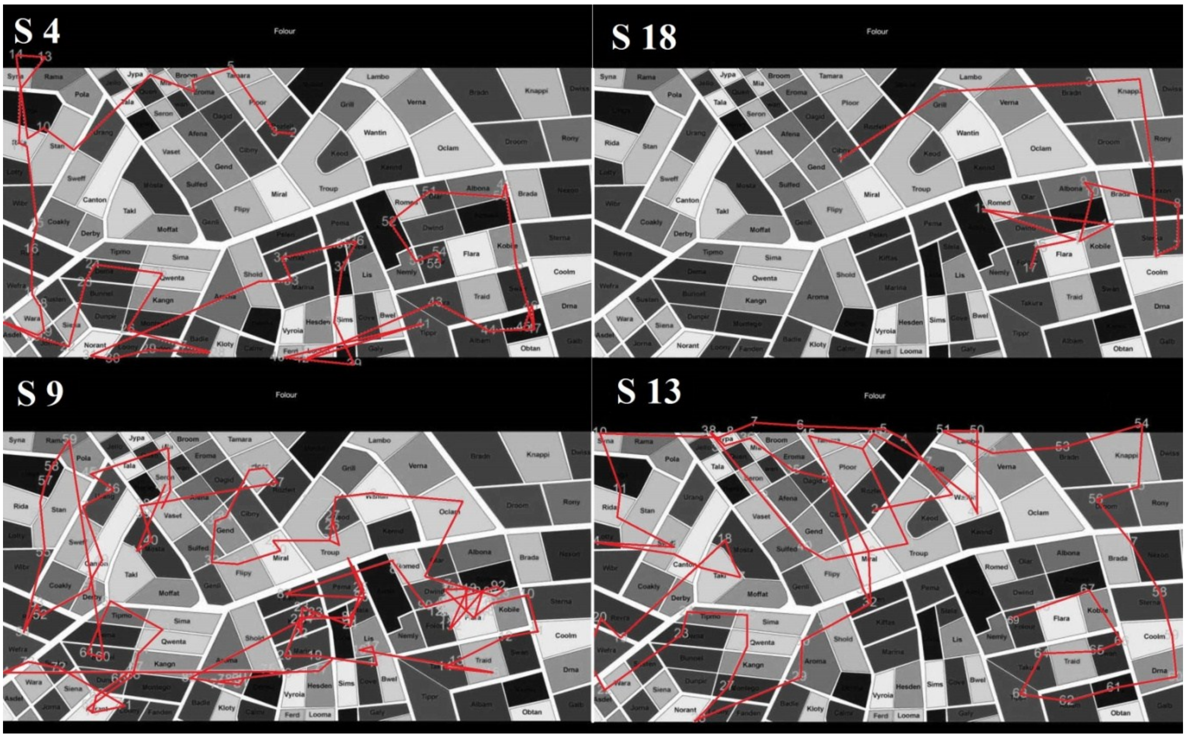


Figure 7.4: An example of four scanpaths for the same stimulus

Moreover, and as shown in Figure 7.5, the search system cannot be described if all the scanpaths are crowded for one stimulus, neither can be distinguished between users groups. This issue makes scanpath analysis harder to achieve at once for the one stimulus. Therefore, data were analyzed in stages using OGAMA and EyePatterns software packages. This usage is described in section 3.1.

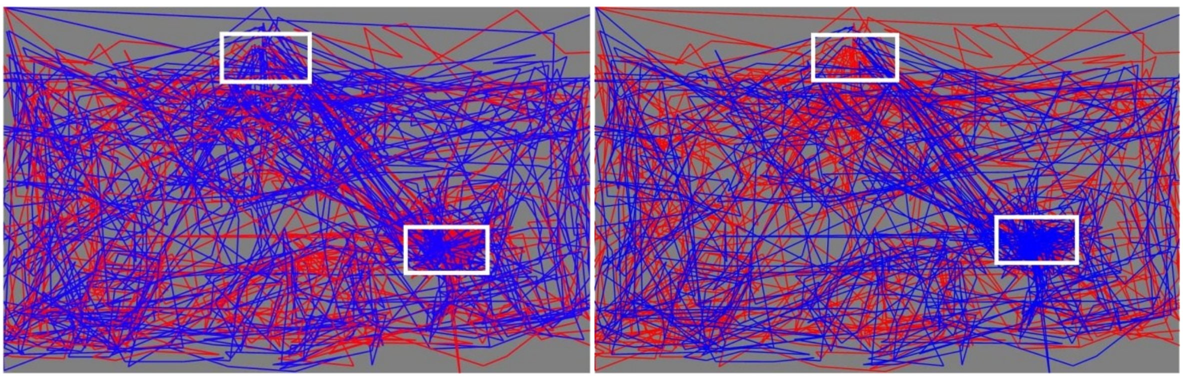


Figure 7.5: An example of scanpath comparison between two groups, (left: experts versus novices, right: females versus males)

7.3.1 Data processing, simplification, and scoring

All the recorded scanpaths were aggregated per stimulus based on the sequence with which the different AOIs were visited. The scanpath strings are determined by OGAMA. Each scanpath was translated into an alignment of letter-codes (from A to P, one for each AOI), based on the order at which they were fixated. These strings, as produced by OGAMA, were then simplified and compressed by EyePatterns as follows:

- LAAAAEEFDDDDDCDCCCCCCCBEEHDCIIIIIIJIIJJJJJMMMMMM→LAEFDCDCBEHDCIJM
- LHHHDDDDDDDECCCIIIIIJJJMMMMMMMNNOPPPPKKKLLLHHGEAEEECCEEEAAEEAAEBDDDDDDCCCCCFFFFEEEEEEGGGGGGHHHHHHLLLLLKKKKKKKKKPPPPPMMMMMMMMMMMMMMM→LHDECIJMNOPKLHGEAECEAEAEBDCFEGHLKPM

Table 7.2: Two scanpath analysis, using rewarding and penalizing

	L	A	E	F	D	C	D	C	B	E	H	D	C	I	J	M
L	+1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
H	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	+1	-1	-1	-1	-1	-1
D	-1	-1	-1	-1	+1	-1	+1	-1	-1	-1	-1	+1	-1	-1	-1	-1
E	-1	-1	+1	-1	-1	-1	-1	-1	-1	+1	-1	-1	-1	-1	-1	-1
C	-1	-1	-1	-1	-1	+1	-1	+1	-1	-1	-1	-1	+1	-1	-1	-1
I	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	+1	-1	-1
J	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	+1	-1
M	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	+1
N	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
O	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
P	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
K	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
L	+1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
H	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	+1	-1	-1	-1	-1	-1
G	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
E	-1	-1	+1	-1	-1	-1	-1	-1	-1	+1	-1	-1	-1	-1	-1	-1
A	-1	+1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
E	-1	-1	+1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
C	-1	-1	-1	-1	-1	+1	-1	+1	-1	-1	-1	-1	+1	-1	-1	-1
E	-1	-1	+1	-1	-1	-1	-1	-1	-1	+1	-1	-1	-1	-1	-1	-1
A	-1	+1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
E	-1	-1	+1	-1	-1	-1	-1	-1	-1	+1	-1	-1	-1	-1	-1	-1
B	-1	-1	-1	-1	-1	-1	-1	-1	+1	-1	-1	-1	-1	-1	-1	-1
D	-1	-1	-1	-1	+1	-1	+1	-1	-1	-1	-1	+1	-1	-1	-1	-1
C	-1	-1	-1	-1	-1	+1	-1	+1	-1	-1	-1	-1	+1	-1	-1	-1
F	-1	-1	-1	+1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
E	-1	-1	+1	-1	-1	-1	-1	-1	-1	+1	-1	-1	-1	-1	-1	-1
G	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
H	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	+1	-1	-1	-1	-1	-1
L	+1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
K	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
P	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
M	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	+1

EyePatterns analyze the compressed strings using either Levenshtein algorithm (string-edit algorithm), or the Needleman-Wunsch algorithm (scoring scheme). Search scanpaths were analyzed using the basic scoring scheme which allow to chose how much to score for each match, gap, and mismatch. This method uses the Needleman-Wunsch algorithm (1970) to determine how similar each pair of sequence is by using the basic scoring scheme each match will score +1, the gap will score 0, and the mismatch will score -1 (West *et al.,* 2006). Table 7.2 shows a scoring scheme between the previous two scanpaths. This method of rewarding each match with +1 and penalizing the mismatch with -1 allowed to calculate the scores of comparing two scanpaths (40 rewarding score in the previous example). These similarities will be visualized in EyePatterns as it will be explained in section 7.3.2. A detailed description of this procedure and its related algorithms can be found in the work of West *et al.* (2006).

7.3.2 Output and visualization

Visualising the search patterns will provide a good comparison between all scanpaths at the same time. This will represent the topological relationships between users' search strategies, which were presented for that task. Because the patterns can highlight groups, or clusters, of sequences that are mathematically similar. When using EyePatterns, this is done by creating a tree that assesses the similarities between sequences based on their scores. The fewer the branches (and thus the nodes between subjects), the similar the scanpath sequences are. Meaning, if two sequences are grouped together connected by one node, they will have the most similar sequence to each other. More nodes between two sequences indicate a lower similarity in the scanpaths and thus the order with which the AOIs are visited (the search strategies). This technique allows searching for clusters within each tree (stimulus). Table 7.3 shows the similarities between scanpaths of the experts and novices when they performed the first stimulus. Table 7.4, A shows the similarities within experts and Table 7.4, B the similarities within novices.

Table 7.3: The scanpath similarity between experts and novices in the first stimulus calculated by OGAMA, experts are placed vertically and novices horizontally

	S16	S17	S18	S19	S20	S21	S22	S23	S24	S25	S26	S27	S28	S29	S30	S31
S1	14%	13%	39%	27%	31%	20%	28%	15%	17%	17%	28%	36%	11%	14%	14%	22%
S2	16%	7%	19%	26%	24%	14%	20%	14%	17%	9%	12%	18%	16%	26%	11%	21%
S3	20%	11%	13%	20%	23%	16%	17%	12%	21%	18%	12%	15%	17%	20%	12%	23%
S4	12%	8%	29%	18%	16%	16%	37%	9%	9%	11%	25%	27%	15%	24%	8%	23%
S5	19%	25%	11%	10%	16%	10%	10%	24%	15%	25%	38%	10%	8%	13%	30%	10%
S6	15%	13%	19%	14%	24%	16%	20%	19%	24%	12%	17%	17%	14%	21%	16%	17%
S7	14%	14%	38%	20%	25%	13%	30%	14%	15%	22%	26%	28%	10%	25%	17%	21%
S8	12%	26%	13%	8%	13%	10%	10%	24%	23%	16%	27%	10%	7%	21%	37%	14%
S9	25%	13%	24%	15%	32%	16%	29%	21%	19%	22%	34%	14%	12%	22%	19%	17%
S10	8%	4%	10%	14%	7%	17%	13%	7%	13%	8%	6%	14%	10%	12%	7%	12%
S11	20%	11%	30%	35%	36%	24%	28%	16%	20%	20%	20%	18%	13%	27%	17%	21%
S12	31%	9%	13%	36%	29%	12%	12%	10%	20%	21%	13%	16%	27%	20%	12%	22%
S13	9%	13%	18%	17%	11%	22%	21%	15%	27%	12%	15%	18%	8%	19%	15%	16%
S14	8%	7%	12%	15%	10%	19%	14%	10%	19%	12%	8%	15%	11%	16%	11%	24%
S15	52%	19%	8%	29%	40%	16%	8%	11%	11%	34%	15%	10%	26%	7%	15%	16%

Table 7.4: The scanpath similarities between experts (A) and novices (B) in the first stimulus presented by OGAMA

S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	
100%	21%	15%	25%	17%	9%	32%	14%	23%	12%	42%	20%	15%	15%	15%	S1
	100%	24%	15%	11%	18%	12%	10%	20%	16%	21%	17%	17%	27%	14%	S2
		100%	12%	12%	14%	12%	11%	14%	11%	13%	30%	26%	20%	26%	S3
			100%	8%	19%	45%	13%	26%	11%	23%	15%	19%	13%	10%	S4
				100%	13%	11%	25%	34%	7%	15%	11%	13%	7%	13%	S5
					100%	19%	16%	29%	12%	18%	16%	19%	18%	9%	S6
						100%	20%	26%	7%	28%	12%	15%	10%	13%	S7
							100%	22%	6%	10%	12%	11%	7%	11%	S8
								100%	8%	21%	18%	19%	10%	18%	S9
									100%	15%	13%	16%	16%	5%	S10
										100%	22%	14%	15%	20%	S11
											100%	18%	17%	36%	S12
												100%	25%	14%	S13
													100%	10%	S14
														100%	S15

A

The scanpath similarity for the experts

S16	S17	S18	S19	S20	S21	S22	S23	S24	S25	S26	S27	S28	S29	S30	S31	
100%	39%	13%	25%	34%	47%	15%	24%	17%	14%	26%	19%	14%	25%	15%	27%	S16
	100%	12%	9%	33%	39%	13%	13%	12%	9%	28%	21%	10%	26%	10%	15%	S17
		100%	9%	8%	13%	8%	6%	24%	19%	24%	19%	9%	6%	16%	15%	S18
			100%	22%	25%	18%	37%	24%	16%	9%	24%	36%	11%	22%	24%	S19
				100%	36%	7%	24%	10%	14%	16%	15%	22%	25%	19%	22%	S20
					100%	19%	30%	20%	11%	27%	21%	22%	20%	24%	16%	S21
						100%	16%	16%	19%	16%	14%	26%	14%	19%	16%	S22
							100%	16%	17%	8%	22%	23%	12%	25%	21%	S23
								100%	19%	18%	15%	14%	8%	21%	12%	S24
									100%	10%	19%	13%	13%	29%	31%	S25
										100%	20%	13%	14%	12%	15%	S26
											100%	15%	8%	16%	15%	S27
												100%	10%	15%	11%	S28
													100%	14%	16%	S29
														100%	21%	S30
															100%	S31

B

The scanpath similarity for novices

The numerical representation is transferred into placements of the similar subject that has the nearest scanpath. Figure 7.6 shows the pattern of the expert group in the first stimuli. The graph shows directly that there is little similarity between subject 8 and subject 10. Meanwhile we need to read 15 ratios from Table 7.4, A to know which is the least similar to the search system of subject 8.

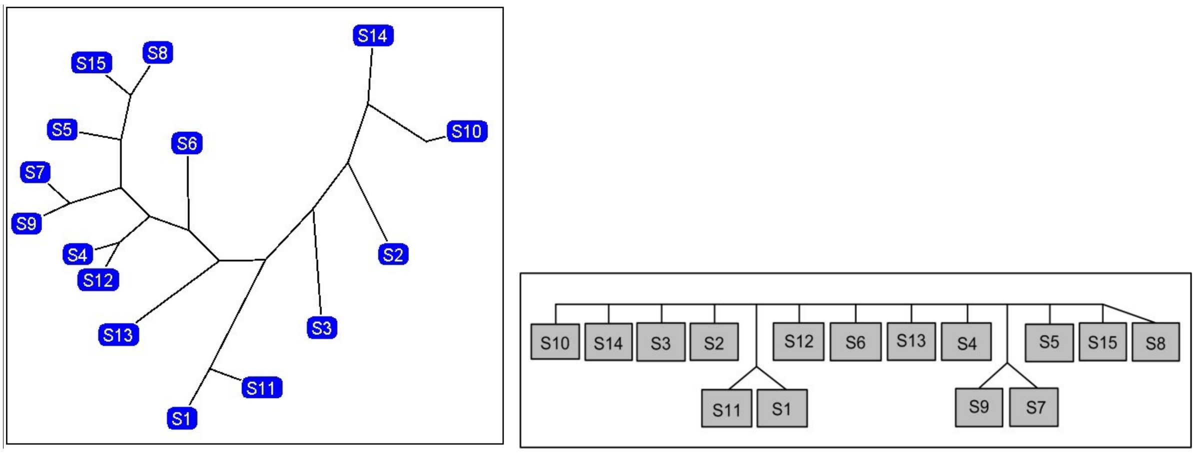


Figure 7.6: Left, visual representation of tree branches showing the degrees of similarities in search systems for expert group generated in EyePatterns, the first stimulus is represented here; right, a redrawn hierarchal pattern by the authors based on the relationships in the original tree

By looking to the patterns (Figure 7.7), they show at least 13 major nodes (stimulus10) and at most 20 major nodes (stimulus12), over the 12 stimuli. This variation indicates the complexity of responses to the task. A few clusters appeared in the patterns as in stimulus2, stimulus3, stimulus4, stimulus6, stimulus7, stimulus8, stimulus9, and stimulus12. The cluster (C) is defined for consecutive subjects in the patterns where it provides:

$$C = nN/nS = 1$$

Where nN is the number of node between the first and last subject in the cluster: nN >3 (10%) of the total subjects' number and nS is the number of subjects in the cluster.

The variety of scanpath can be linked mainly to the complexity of the presented task as a subject has to scan 108 labels to find the target label. This result is compatible with Brandt and Strak (1997) as variations in the stimulus (target name and placement) presented a considerable variation in scanpath patterns.

7.3.2.1 Expertise variations

Two user groups could be distinguished: experts and novices. Figure7.7 illustrates the results related to the 12 different target labels; Experts are presented in grey and novices are presented in white. For most of the stimuli, there is no outstanding cluster between experts and novices apart from some clusters. Different clusters can be found for the experts' group in stimulus6 and stimulus12. And some clusters are found for novices' group in stimulus6, stimulus8, stimulus9, stimulus12. This lack of clusters in the 12 patterns could have occurred because none of the subject knows the map stimuli and their design in advance, and most of the subjects developed their search system between the first and the last map. This development is noticeable in the graphs, as the pattern, the number of nodes, and the typological relationship changed over stimuli. Where the major nodes, the secondary nodes, and the other levels of nodes' relationships appear differently in each patterns. This reflects different scanpath for each stimulus. In addition to that, the task of finding label does not consider as a typical task where the education and training influence vastly the users' responses. Unlike reading contour lines, understanding chart maps, and decoding satellites images.

7.3.2.2 Gender variations

Concerning the gender influence on search system (females and males), the pattern are distinguished and illustrated in Figure 7.7 and Figure 7.8 (females in red and males in blue). Females' subgroup shows some clusters like in stimulus4, stimulus8, stimulus7, stimulus12. And males' subgroup shows some clusters too, like in stimulus2, stimulus4, stimulus7, stimulus9 stimulus12. The gender characteristic did not influence the scanpath in this application. Since the gender effect on the search task did not show a pattern with which can be described for all the clusters. Users' characteristic of gender randomly influenced the clustering in the pattern as no evident numerical or topological relationships appeared along these 12 patterns.

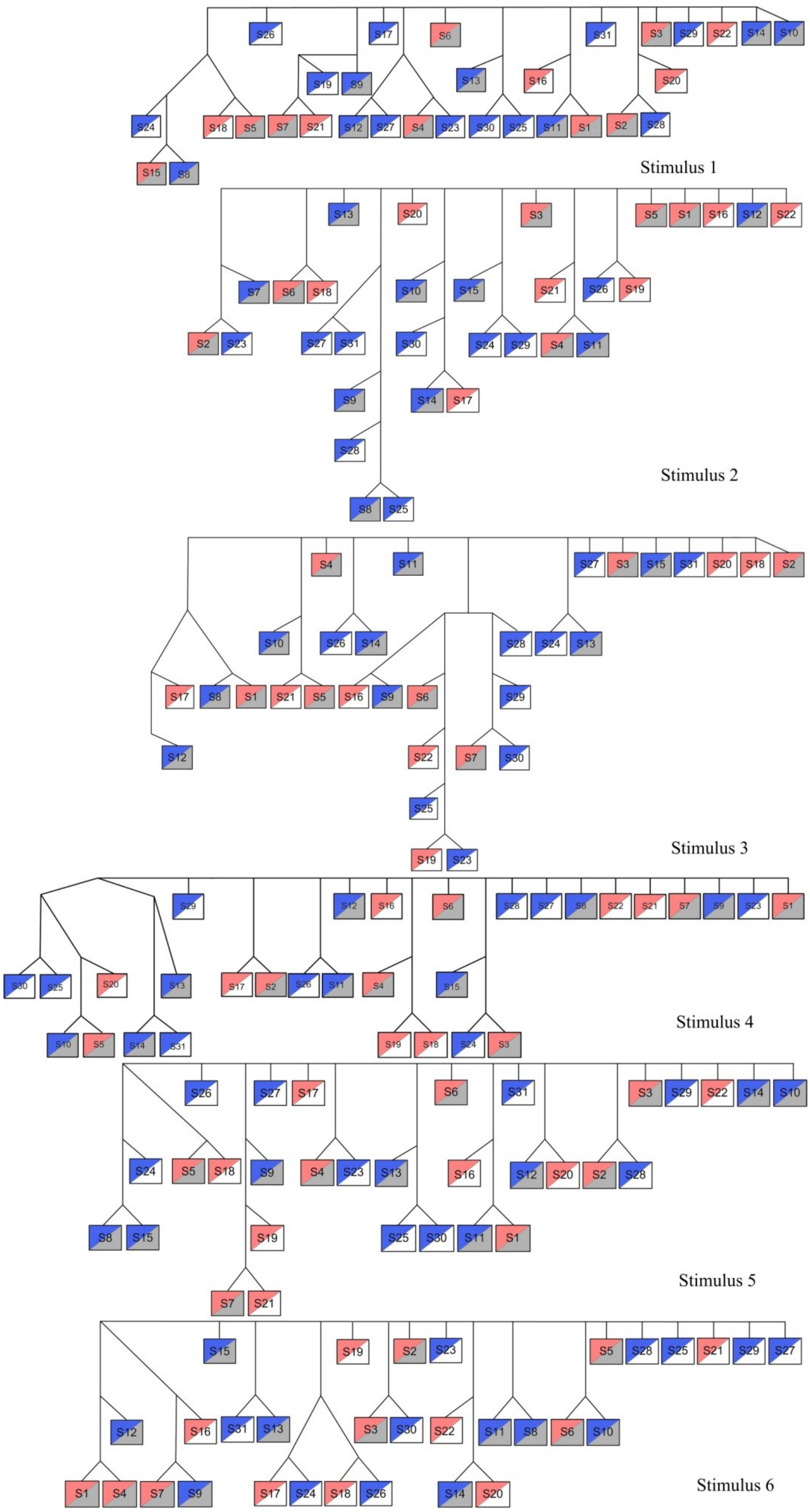


Figure 7.7: Users' performance on 6 stimuli (1-6) considering the expertise and Gender variations, experts (grey) versus novices (white) and females (red) versus males (blue)

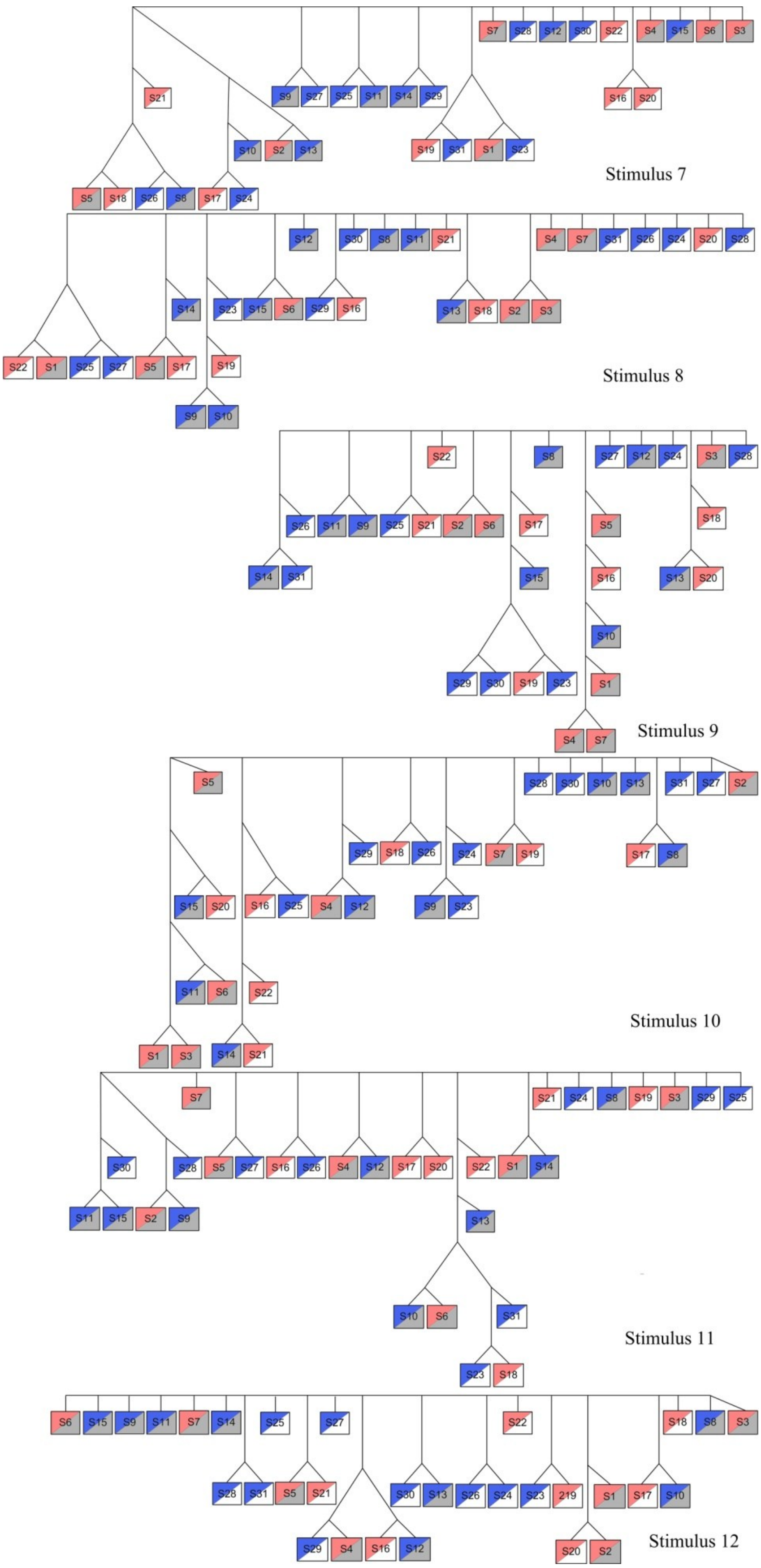


Figure 7.8: Users' performance on 6 stimuli (7-12) considering the expertise and Gender variations, experts (grey) versus novices (white) and females (red) versus males (blue)

7.3.3 Clusters' description

Since the eye patterns do not provide any factual dimensional scaling of the stimuli, it was important to link the mathematical clusters [C = nN/ nS =1] of the scanpaths to the original scanpaths. Figure 7.10 illustrates the pattern of scanpaths made at stimulus 9. The pattern introduces a cluster of search systems made by males. This cluster contains subjects (25, 9, 11, 26, 31, and 14). There is six nodes between subject 25, and subject 14 (25 is the first subject in the cluster and 14 is the last subject in the cluster). The number of nodes reflects the dissimilarity gap between the subjects. So we shall expect 6 variations of the scanned AOIs. Each of which has at least on gap of the scanned AOIs and at most six gaps of them. To clarify this issue, the scanpaths shall be monitored back in OGAMA software. Figure 7.9 illustrates 6 scanpaths of the mentioned cluster. It presents that each string has one difference to the following string related to their appearance in AOIs. The pattern shows different level of tree branches. The described pattern was located at the second and third level of the tree branches. Each string has always, at least one difference between the previous and the next in the described pattern. The relationship between scanpaths can also be derived by their locations in different levels of the tree, where consecutive subjects at the same level are more consistent (S31 and S 14) than the other consecutive string (S26).

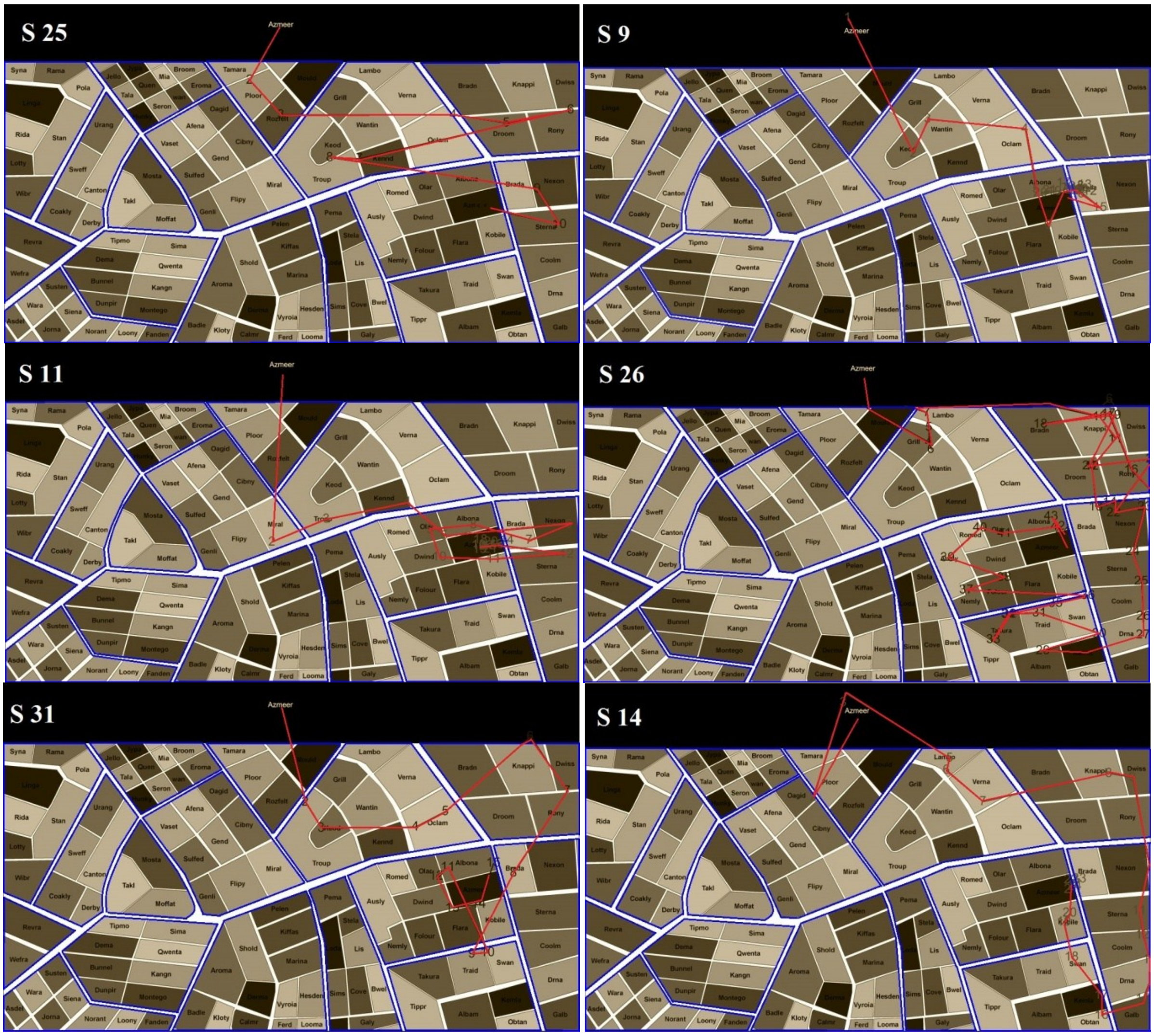


Figure 7.9: Scanpath patterns in OGAMA, male cluster in stimuli 9

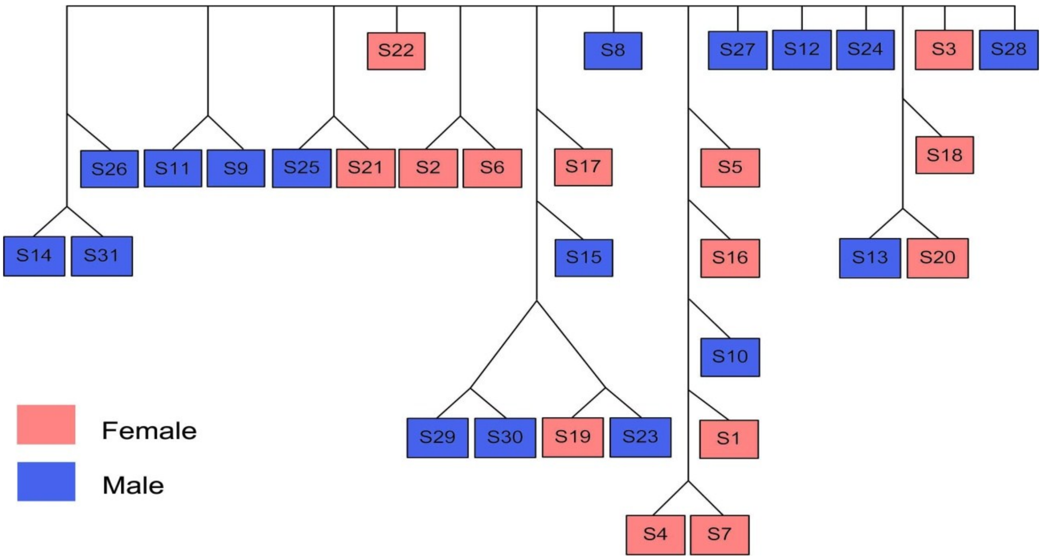


Figure 7.10: Scanpaths pattern in EyePatterns. Users' performance on stimulus 9, gender variation

7.4 Conclusion

The study described the use of multi-package open source softwares to analyze a search task for a target label on maps. The search scanpaths were analysed using OGAMA and EyePatterns softwares. The results categorized the scanpaths based on AOIs' shapes and dimensions. The patterns were not consistent over the 12 stimuli. Under the circumstances of the presented study, this attempt to link the subject characteristic with the motivational variable has failed to show consistent clusters. Although some minor clusters based on expertise and gender appeared, the results here gave an insight about the variability of searching strategies over map face.

Visualizing the scanpath with EyePatterns provided the ability to compare and distinguish patterns and group of users. The study showed that searching for label target among less number of labels and less complex background is expected to introduce more unified patterns.

Further research work shall be conducted to determine the patterns in which users searching system can be identified by using similarities analysis between scanpaths in Visual Analytics Toolkit (Andrienko *et al.,* 2012). Where data can be treated using methods like spatial generalization, adjustment of time reference, and spatio- temporal aggregation.

7.5 References

Andrienko, G., Andrienko, N., Burch, M., & Weiskopf, D. 2012. Visual Analytics Methodology for Eye Movement Studies. IEEE Transaction on Visualization and computer Graphics, 18(2), pp.2889-2898.

Andrienko, G., Andrienko, N., Rinzivillo, S., Nanni, M., Pedreschi, D., & Giannotti, F. 2009. Interactive Visual Clustering of Large Collections of Trajectories. IEEE VAST 2009 .pp. 3-10, Atlanta City, NJ.

Augustyniak, A. & Mikrut, Z. 2006. Complete scanpaths analysis toolbox. In 28th IEEE EMBS 2006 pp.5137-5140, NY.

Augustyniak, P. 2006. Monitoring the progress of fast reading training with use of the eyetracker and scanpath statistics. Journal of Medical Informatics & Technologies, 10, pp.153-162.

Augustyniak, P. 2003. How a human perceives the electrocardiogram - the pursuit of information distribution through scanpath analysis. Computers in Cardiology , 30, pp.601−604.

Brandt, S. A. & Strak, L. W. 1997. Spontaneous Eye Movements During Visual Imagery Reflecte the Content of Visual Scene. Journal of Cognitive Neuroscience, 9(1), pp.27-38.

Byrne, M. D., Anderson, J. R., Douglass, S. & Matessa, M., 1999. Eye Tracking the Visual Search of Click-Down Menus. In the SIGCHI conference on Human Factors in Computing systems, pp.402-409, NY.

Cleveland, W.S. & McGill, R. 1983. A Color-Caused Optical Illusion on a Statistical Graph. The American Statistician, 37(2), pp 101-105.

Duchowski, A. T. 2007. Eye Tracking Methodology, Theory and Practice. (2nd ed.). Springer. London.

Duchowski, A. T. 2002. A Breadth- first Survey of eye tracking Applications. Behaviour Research Methods Instructions & Computing, 34(4), pp.455-470.

Gellatly, I. R. 1996. Conscientiousness and Task Performance; Test of Cognitive Process model. Journal of Applied Psychology, 81(5), pp.474-482.

Haklay, M. M. & Nivala, A. M. 2010. 'User-centred Desig', In Haklay, M. M (eds.), Interacting with Geospatial Technologies,chapter 5, pp. 91-106. John Wiley& Sons. Chichester.

Hornof, A. J. & Halverson, T. 2003. Cognitive Systems and Eye Movements for Searching Hierarchical Computer Display. In the SIGCHI Conference on Human Factors in Computing Systems, pp.249-256, NY.

Ooms, K., Andrienko, G., Andrienko, N., De Maeyer, P. & Fack, V. 2012. Analysing the spatial dimension of eye movement date using a visual analytic approach, Expert Systems with Applications, 39 (1), pp.1324-1332.

Phillips, R.J., Noyes, E. & Audley, R.J. 1978. Searching for names on maps. The Cartographic Journal, 15(2), pp.72-77.

Phillips, R. J. 1981. Search for Target in a Random Arrangement of Names: an Eye Fixation Analysis. Canadian Journal of Psychology, 35(4), pp.330-346.

Lleras, A. & Von Mühlenen, A. 2003. Spatial Context and Top-Down Strategies in Visual Search, Spatial Vision, 17(4). pp. 465-482.

Maltin, M. W. 2003. Cognition. (5th ed.). John Wiley &Sons, NJ.

Nielsen, J. 1993. Usability Engineering. Academia Press, San Francisco.

Rayner, K. 1998. Eye Movement in Reading and Information Processing: 20 years of Research. Psychological Bulletin, 124(3), pp.372-422.

Rubin, J. & Chisnell, D. 2008. Hand Book of Usability Testing: How to Plan, Design, and Conduct Effective Tests. 2nd edition, John Wiley &Sons, Indianapolis.

Strak, L. W. & Ellis. S. R. 1981. Scanpaths revisited: cognitive models direct active looking. In Fisher, D.F *et al.* (eds.), Eye movements: Cognition and visual perception (pp.193-226), Lawrence Erlbaum Associates, Hillsdale, NJ.

West, J. M., Haake, A. H., Rozanski, E. P. & Karn, K. S. 2006. EyePatterns: Software for Identifying Patterns and Similarities Across Fixation Sequences. In the Symposium on Eye Tracking and application, pp.27-29, San Diego, California.

Wolfe, J. M. 1994. Guided search 2.0: A revised model of visual search. Psychonomic Bulletin & Review, 1(2), pp.202-238.

Yang, G-Z., Marco, L. D., Hu, X. P. & Rowe, A. 2002. Visual search: psychological models and practical applications. Image and Vision Computing, 20(4), pp.291-305.

8

Background and Foreground Interaction, and the Influence of Complementary Colours on Search Tasks

Modified from: Deeb, R., Ooms, K., Van Etvelde, V. & De Maeyer, P. In press, Background and foreground interaction: complementary colours' influence on search task. Color Research and Applications, DOI: 10.1002/col.2192

Preface

Labels are indispensable visual (communication) elements that completely deliver the geospatial message of maps. The objective of this study is to investigate the impact of complementary colors between the map's background and text on the readability of cartographic texts and thus on the efficiency of the map user's search task. This is compared with the use of the 'traditional' black labels on the corresponding coloured backgrounds. Furthermore, a number of user characteristics, such as gender and expertise, are taken into account as well. The users' eye movements were registered to study their attentive behavior during the visual search task. In addition to the complement of the colour's hue, the analyses were based on the difference in luminance, which could also affect the labels' readability. The difference between the black and coloured label design was significantly different versus the eye-tracking metrics. A correlation was found between the colour difference and reaction time measurement and between the luminance difference and fixation duration.

8.1 Introduction

In the past, researchers have investigated how a visual search is guided by the visual variables of size, shape, orientation, colour (hue), value and texture (Bertin, 1967; Wolfe, 1994; Foley *et al.,* 2007). These visual variables are essential for distinguishing elements in the visual world, often depicted on maps. They are implemented in different stimuli to assess their influence on target searching tasks (Wolfe, 1994; Foley *et al.,* 2007; Benjamins *et al.,* 2009; Nachmias, 2011; and others).The stimuli themselves define the way in which the target is interpreted and perceived and thus play a basic role in reaction time studies (Strak and Ellis, 19881; Wolfe *et al.,* 1989; Brandt and Strak, 1997). These variables were studied individually and were also combined. Wolf *et al.* (1989), for example, analyzed the shape, size and colour search individually as well as in a dual and triple conjunction in relation to the users' search behavior. They reported that triple conjunctions are more efficient than standard conjunctions because three parallel processes can guide the search more efficiently (as opposed to two). The perception of size and shape - more specifically the difference between size and shape - was analyzed by Nachmias (2011) who found that the discrimination between the size and shape can be enhanced by the method of presentation. Colour and orientation were examined by Anderson *et al.* (2013), and they concluded that these visual variables operate in different ways to guide the search. It is noteworthy that an efficient search can be provided by avoiding small differences between targets and by presenting independent distractions (Benjamins *et al.,* 2009).

The visual variables have different properties and capacities for a given type of information (Bertin, 1967). Deeb *et al.* (2012) explained these properties for which the variables are applied on the typographic design. Thorough studies discussed how these variables influence the typographic perception and affect the efficiency of reading plain text. Chan and Lee (2005) tested how font type (shape), character size and line spacing affect the reading speed of Chinese scripts and concluded that only character size had a significant influence on the comprehension score. Bernard *et al.* (2003) compared the effects of text size and shape on the readability of computer displays. They tested Times New Roman and Arial text and concluded that the text size and shape had significant effects on the readability of both designs. Both text reading tasks and visual searching tasks were employed by Huang *et al.* (2009) to reach an empirical answer regarding the optimal font size for Chinese characters. The variation of the typographic shape between the upper and lower case was studied by Arditi and Cho (2007). They concluded that the upper case is more legible than the lower- and mixed-case style. Garcia and Caldera (1996) tried to find the best combination of background and foreground colours that maximize the readability of on-screen displays. They investigated blue, yellow, white, grey and black. In addition to the colour combination, they studied the shape (typeface) and size of the typeface. The tested colour combinations showed significant effects for several combinations (black and grey), but no significant effects on the reaction time for the colour and shape combinations. The mentioned studies attempted to explain the plain text reading and defined several rules for an effective typographic design. Few studies tackled the reading of cartographic text where the legibility of the labels is ruled by:

- The variability in which the text is depicted (i.e., how the visual variables were implemented in the design);
- The designed contrast of the text;
- The distribution of text over the map among other features; and
- The display resolution and medium.

Colour is a vital variable in map design because coloured maps are more preferred than monochrome maps (Brewer, 1997). Cartographers stressed that colour is worth the extra effort and expense; it adds improved map making because it permits a greater accuracy in map reading. To emphasize visual differences, Brewer (1997) suggested that the contrast can be enhanced if saturation differences were used. In addition, the diverging colour schema proved to obtain more accuracy in users' responses for the design of choropleth maps (Brewer *et al.,* 1996). Moreover, Slocum *et al.* (2005) discussed the influence of surrounding colours and stressed that induction causes the colour of an area to shift towards the complementary colour of the surrounding colour. However, the effect of background colours on other map elements in the foreground was not defined before. It is noteworthy that the effect of background colours on the colour search was analyzed by Rosenholtz *et al.* (2004), who confirmed that asymmetries in the colour search depended on the relationship between the search target's colour and background colour. Additionally, De Vries *et al.* (2013) studied different backgrounds and concluded that darker backgrounds require a longer search time for the same target. Moreover, the search time was the longest for small target-background differences. Moreover, Carter and Huertas (2010) investigated the effect of an ultra-large colour difference on enhancing the conspicuousness and discriminability of small substances. They stressed the effect of the background luminance on the target.

This chapter presents an eye-tracking experiment that was conducted to answer three questions: first, what is the influence of the complementary colours (background-label) on the users' search efficiency; second, is this further influenced by the user's characteristics (gender and expertise); and finally, are the users' preference and search efficiency linked to each other? Furthermore, the findings are compared with the more 'traditionally' used black labels (on the same background) as a validation. The study design is explained in detail in the following section.

8.2 Study design

8.2.1 Apparatus

A controlled experiment was conducted in the eye-tracking laboratory of the Department of Geography, Ghent University, Belgium. The laboratory is equipped with SMI RED 250 with a sampling rate of 120 Hz (SensoMotoric Instruments GmbH, Germany). The stimuli were displayed on a 22-inch DELL screen (1680×1050 pixels; refresh rate: 60 Hz). Participants were seated comfortably at a 50-cm viewing distance from the monitor, and the monitor's height was adjusted in accordance with the participants' height to perform the best calibration.

8.2.2 Participants

Because user characteristics are crucial to capture and need to be taken into account during the experimental design (Haklay and Nivala, 2010; Nielsen, 1993; Rubin and Chisnell, 2008; Duchowski, 2007; and others), the study has both a between- and within-user design. In total, 31 participants volunteered to do the test. They had a 20/20 acuity or wore correcting lenses. Olson and Brewer (1997) clarified the influences of colour-vision impairment on map readability. As a consequence, to avoid biased results, none of the participants had anomalous colour vision. Two groups of participants were distinguished. The first group consisted of 15 experts who were in daily contact with cartographic materials and obtained at least a Master's degree in Geography or Geomatics. Out of the 15 participants, 7 were female and 8 were male. The second group of participants consisted of 16 novices. The novices were first-year Bachelor students who had just started their education in

Geography or Archeology. Thus, they did not receive any cartographic training beforehand. The novice group consisted of 7 females and 9 males. The average age was 25.9 years for experts, 19.1 years for novices, 23.3 years for females and 23.9 years for males.

8.2.3 Map design

A fictive land use map was designed for this study. It consisted of polygons that were filled in basic colours as explained in Table 8.1 and Figure 8.1. In total, 24 maps were constructed, which have the same basic background design. However, a variation in the labels' colour (on the foreground) was presented in each map. The colours (background and foreground) were chosen so that the colour difference would be the highest: value = 100%. Complementary colours (hue) were used for the background and foreground (i.e., text) to enhance the conspicuousness of the foreground feature. In addition, a second set of maps with black labels was constructed to validate the obtained results. Twelve maps were thus populated with coloured labels and 12 corresponding maps were populated with black labels. The selection of the 12 colours to be used for the coloured labels is illustrated in Figure 8.1. To obtain an accurate description of how the colours were actually presented to the users, they were also measured on the screen on which all stimuli were displayed. These measurements were conducted using an i1Pro from Gretag Macbeth/X-Rite in combination with the Measure Tool software (see third and fourth column, Table 8.1). The measured colours were not identical to the 'theoretical' colours that were defined when designing the maps. Additionally, the colour space that describes the measured colours deviated from the normal colour space because the measurements reached luminance values higher than 100. This distortion was mainly caused by the characteristics of the screen on which the stimuli were presented (in comparison to normal daylight conditions). To validate these measurements, an ICC profile of the screen was created (using the i1Pro in combination with the ProfileMaker software). The resulting profile confirmed the measured values.

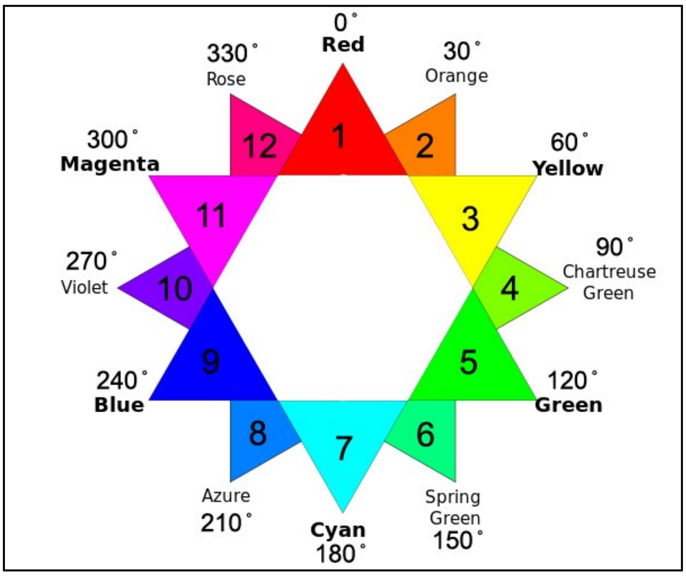


Figure 8.1: The twelve tested colour with their common names

Target labels (for the search task, see next section) were placed on the map face so that its colour would correspond to the complement of the background colour (H + 180°). For each coloured target, a corresponding black target was designed. Figure 8.2 shows the distribution of the 12 targets over the corresponding 12 background colours. Because the design used one of the 12 colours for labels in the foreground, it was necessary to exclude that colour from the background. This was performed simultaneously for each pair of corresponding maps (with the coloured and black label). To avoid any influence on the working memory, the letters of the target labels were different within each pair. However, they provided a similar cap height, x-height, counter space, loop and bowl (Arditi and Cho, 2005; Arditi and Cho, 2007) Furthermore, the number of letters in each corresponding target label was preserved to ensure that targets had the same length for a valid comparison (Phillips, 1981).

Table 8.1: The used colour in the stimuli presented in HSV and RGB when the maps were designed and the measured CIE and XYZ systems when the maps were displayed

Colour system	**Design conditions**						**Display conditions**					
	HSV			**RGB**			**CIE**			**XYZ**		
Colour	H°	S%	V%	R	G	B	L*	a*	b*	X	Y	Z
1	0,	100	100	255	0	0	69.9	95.7	77.5	76.09	40.18	4.617
2	30	100	100	255	128	0	86.0	48.6	79.7	88.28	67.98	11.92
3	60	100	100	255	255	0	121.8	-24.3	101.1	140.21	167.63	34.10
4	90	100	100	128	255	0	115.3	-90.6	90.3	81.46	145.01	33.79
5	120	100	100	0	255	0	112.3	-111.5	86.9	65.28	135.30	32.49
6	150	100	100	0	255	128	111.2	-99.6	40.6	68.50	131.85	76.55
7	180	100	100	0	255	255	116.5	-64.8	-39.4	98.45	149.03	257.74
8	210	100	100	0	128	255	70.6	20.4	-109.4	46.27	41.60	232.25
9	240	100	100	0	0	255	45.6	87.8	-148.7	33.45	14.97	222.16
10	270	100	100	128	0	255	55.5	94.3	-132.2	49.45	23.41	223.65
11	300	100	100	255	0	255	71.7	101.5	-6.3	83.62	43.21	52.41
12	330	100	100	255	0	128	79.1	114.9	-92.2	109.63	55.10	225.46
Black	0	0	0	0	0	0	1.5	0.8	-5	0	0.2	0.2

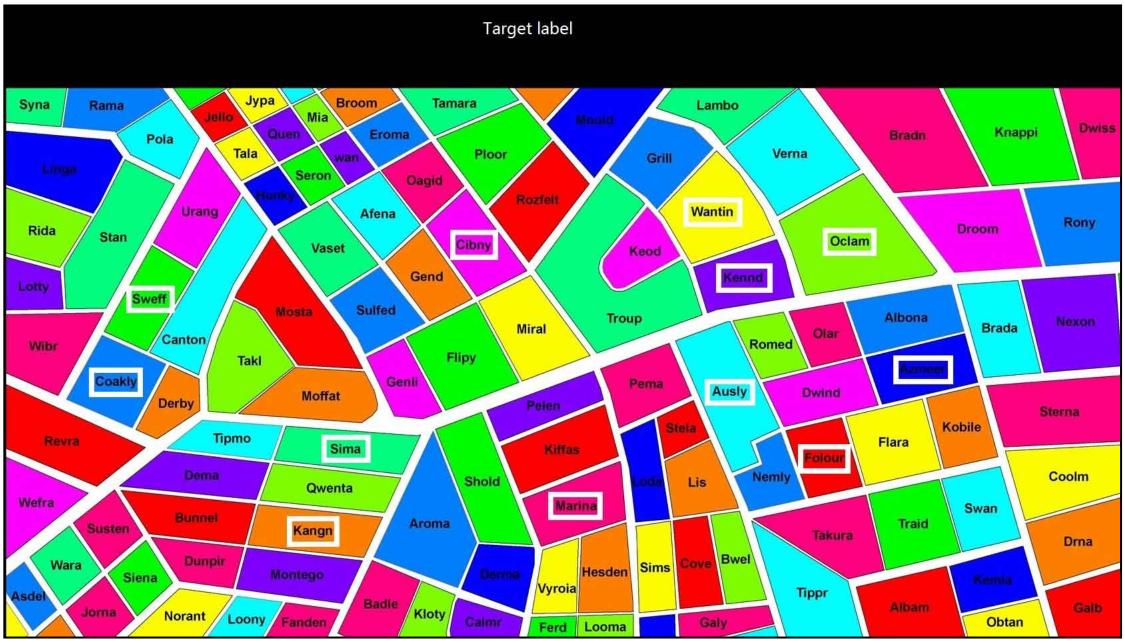


Figure 8.2: Target label distribution of the map face

8.2.4 Stimuli, task and procedure

In total, the 24 map stimuli were presented to all participants in two phases of 12 maps each. Before starting a phase, a calibration of the eye-tracking device was executed. The two label designs (coloured versus black) were presented alternatively and randomly in the trail to equalize the influence of participants' fatigue on their responses. This randomness was also important to prevent biases because of the learning effect. For each display, participants were asked to locate the target label that was displayed on the top of the map. Thus, the eye movement data were recorded starting from the target label over the map and ending with the target on the map (see Figure 8.2). When participants found the target, they were asked to look at it for two seconds and then the next map automatically appeared. Before the actual start of the test, verbal and written instructions were given. Then, a demo map was presented to the participant. The purpose of the demo map was to introduce the stimuli environment to the participant and prevent biases in the first map of the trial. A post experiment questionnaire was presented at the end of the trail to track participants' characteristics, such as age, gender, expertise level and their observations about the stimuli. The stimuli design is illustrated in Figure 8.3.

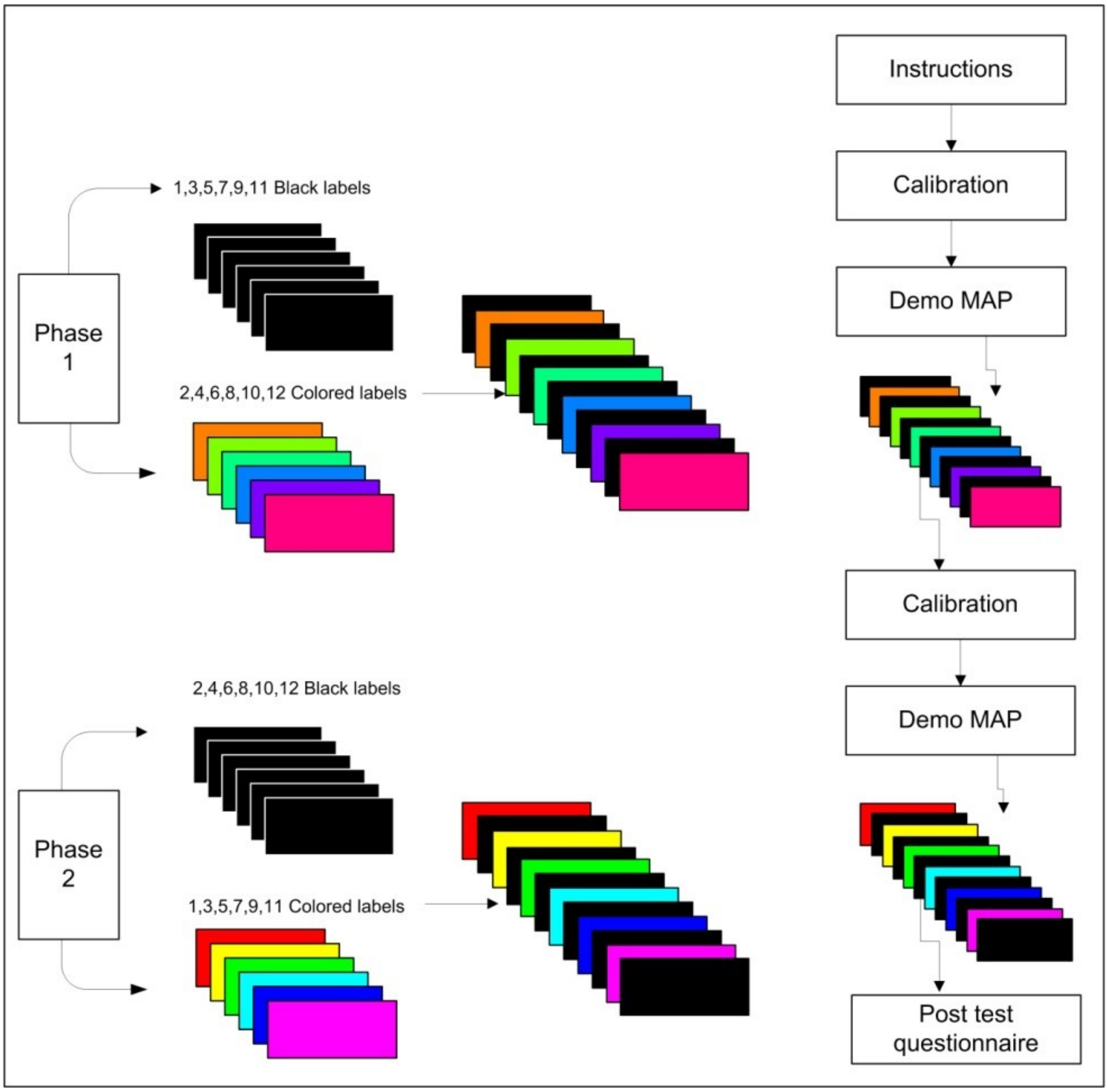


Figure 8.3: the stimuli design with an example of map order

8.2.5 Data and recordings

The users' attentive behavior on the stimuli was registered with an SMI RED250 eye-tracking device, which registers the position of the participants gaze at 120 Hz. The last two seconds of all records were subtracted from each trial because these seconds were implemented in the users' task design (as explained in section 8. 2.4) and were thus not related to the actual responses. Furthermore, several individual records had to be excluded from the data analysis because their measurements showed errors. These errors mainly occurred because the eye tracker lost track of the participant's eyes. This issue is a two-fold problem. First, several participants had corrected vision so they had to wear lenses to achieve good calibration. However, when their lenses became dry, the tracker missed the track of their eyes and recorded extra fixations that could be misinterpreted if they were taken into account while analyzing the data. Second, when participants reached the end of the stimuli, the duration of the trial exceeded the mean + 2 × standard deviations. This could indicate fatigue. Therefore, these errors were excluded from the data.

8.3 Results

8.3.1 Response time measurements

The response time measurement reflects how fast users can perform the task and their efficiency (Nielsen, 1993). A one-sample t-test was executed to check the difference in the users' efficiency regarding the 12 map designs and a comparison between black and coloured labels. The test showed that the only significant difference between participants' efficiency for the black and coloured label was detected in colour12 (H=330̊, S=100%, V=100%) and colour6 (H=150̊, S=100%, V=100%) in the background (colour12: mean=38.613 s; black label mean=14.5 s; F= 8.314, P=0.006). For the complementary colour combination, it was too difficult to read the labels, whereas a black label was more legible over the used background. It is worth mentioning that several coloured labels and their background combinations recorded a higher efficiency than black labels, such as labels in colour2, colour7 and colour11, as indicated in Table 8.2.

Table 8.2: A comparison of the users' responses (s) between black and coloured labels (M= Mean, SD= Standard Deviation)

Map Number	**Black**		**Colour**		***F***	***P***
	M	***SD***	***M***	***SD***		
1	15.932	10.603	20.955	15.622	2.077	0.155
2	20.252	21.420	13.672	10.090	2.217	0.142
3	18.075	13.104	17.174	13.829	0.069	0.793
4	14.972	22.713	17.785	14.344	0.319	0.574
5	13.814	14.905	18.299	21.648	0.089	0.766
6	23.342	198.80	32.562	38.221	1.328	0.254
7	20.653	14.476	14.876	13.489	2.476	0.122
8	14.511	12.934	14.822	13.136	0.009	0.927
9	13.501	11.750	18.277	13.847	2.144	0.148
10	16.589	12.404	20.589	12.404	1.300	0.259
11	26.218	25.308	16.940	12.609	0.179	0.674
12	14.560	10.138	35.918	38.613	8.314	**0.006**

The differences in the users' reaction time between both black label designs and coloured label designs were analyzed by using the multivariate analysis of variance (MANOVA), which showed a significant difference between the users' efficiency over the stimuli (F= 4.519, P=0.000). In addition, the users' performance according to their characteristics was included in the statistical model: gender (female versus male) and expertise (expert versus novice). Table 8.3 shows that none of the expertise groups, gender groups or their interactions with the map design and with each other showed a significant influence on the users' reaction time. Only the map design showed a significant influence on the users' reaction time and thus their efficiency on locating target labels.

When looking at the details of the pairwise comparisons by the Tukey post hoc test, a significant difference between the map designs was found in relation to the map 12 colour label (colour6 as the labels' background) and map 10 with black labels (colour4 as the labels' background). The differences for the mean reaction time with all other maps were positive differences. Although in this case the Tukey test is not very powerful because of the many pairwise comparisons, it still gives a good impression on which designs are the least efficient for the user. Consequently, it can be concluded that both designs (label colour12 – background colour6; label colour black – background colour4) are the least efficient designs.

8.3.2 Fixation duration

A longer fixation duration could indicate difficulty in extracting information (Rubin and Chisnell, 2008). How efficiently both designs were interpreted by the users was evaluated and compared by applying the MANOVA test (across the eye-tracking metrics, including the obtained fixation durations), see Table 8.3. This test indicates that the interaction between the expertise group, gender group and map design is not significant. The only significant difference was located on the map design (F= 2.756, P=0.000).

A more detailed pairwise comparison regarding the map designs (or map numbers) was performed using a Tukey post hoc test. The only significant difference could be located for map 12 (with coloured labels) and map 10 (with black labels), which confirms the findings of the reaction time measurements. In addition, these two map designs are linked with significantly longer fixations, which can indicate a higher cognitive load to process the stimuli (and more exactly the label which has to be located) in comparison to the other 22 map designs.

Table 8.3: The statistical model of the MANOVA test regarding the map design, users' expertise and users' gender for reaction time measurements(s), fixation duration (s) and fixation count (Fix/s)

Source	**df**	**Reaction Time(s)**		**Fixation Duration (s)**		**Fixation count**	
		F	***P***	***F***	***P***	***F***	***P***
Corrected Model	117	2.079	0.000	2.240	0.000	1.518	0.001
Intercept	1	354.591	0.000	535.231	0.000	3343.520	0.000
Map number	**23**	**4.519**	**0.000**	**2.756**	**0.000**	**1.930**	**0.000**
Expertise	1	1.361	0.244	0.055	0.814	0.185	0.667
Gender	1	0.996	0.370	0.037	0.964	0.290	0.748
Map number * Expertise	23	1.000	0.463	0.105	1.000	0.878	0.629
Expertise * Gender	1	0.009	0.925	1.024	0.312	0.082	0.775
Map number * Gender	44	1.037	0.410	0.244	1.000	0.679	0.944
Map number * Expertise *	23	0.605	0.927	1.033	0.420	0.706	0.842

8.3.3 Fixation count (fix/s)

The structure of the stimuli could influence the users' search behavior (Phillips, 1981; Ooms, 2012).Two main factors can control the fixation registrations: search strategies, which lead and define the users' scanpath (Arditi and Cho, 2005), and the cognition and perception difficulty of each stimulus (Hegarty *et al.*, 2010), ruled by the legibility of the studied colour and its complement. A higher fixation count could indicate a lower cognitive load (Rubin, J. and Chisnell, 2008; Harrower, 2007) to read the labels' details. Furthermore, it can show that the colour combination (label – background) is less or more difficult to read in comparison to other combinations.

To obtain a thorough overview of both designs (black versus colour) in relation to expertise and gender differences, the mean fixation counts per second were analyzed using the MANOVA test (see Table 8.3). The analysis showed no interaction between the three variables. The only significant difference was found in the map designs (F=1.930, P=0.000).

For a more detailed result over the 24 label designs, the Tukey test was conducted for pairwise comparisons between the map designs. The number of fixations per second was significantly lower for map 12 (with coloured labels) and map 10 (with black labels) in 20 out of 24 pairwise comparisons.

8.3.4 Colour difference and luminance contrast

The discrimination of targets is not only influenced by the hue of a colour(or the use of complementary colours) but also by the level of colour difference (ΔE*ab) between the target and background (Williams, 1967; Rosenholtz , 2004; Carte and Huertas, 2010; De Vries *et al.,* 2013) and the luminance contrast between both the target and background as well (Shlaer, 1937; Williams, 1967; Johnson, 1995). Therefore, the correlation between the colour difference (foreground-background) and the obtained measurements (reaction times, fixation duration and fixation count) were verified as well. In these calculations, the measured colours are included because they give a better approximation of how the participants observed the colours (instead of the theoretical values that were used to design the maps). The colour difference between the foreground and background was calculated as (Werman, 2012):

$$\Delta E^*_{ab=} \{(\Delta L^*)^2+(\Delta a^*)^2+(\Delta b^*)^2\}^{1/2}$$

where:

- $\Delta L^*= L_{foreground}^* - L_{background}^*$;
- $\Delta a^*= a_{foreground}^* - a_{background}^*$;
- $\Delta b^*= b_{foreground}^* - b_{background}^*$.

The results are visualized in Figure 8.4: a scatter plot of the calculated colour difference and the measured reaction times. The results, under our experimental conditions, showed a negative correlation between the variation in colour difference and user efficiency (more colour difference enhanced performance by reducing reaction time), when the reaction time measurements of both the coloured label designs and black label designs were combined and statistically analyzed (P= 0.045).

The regression model is:

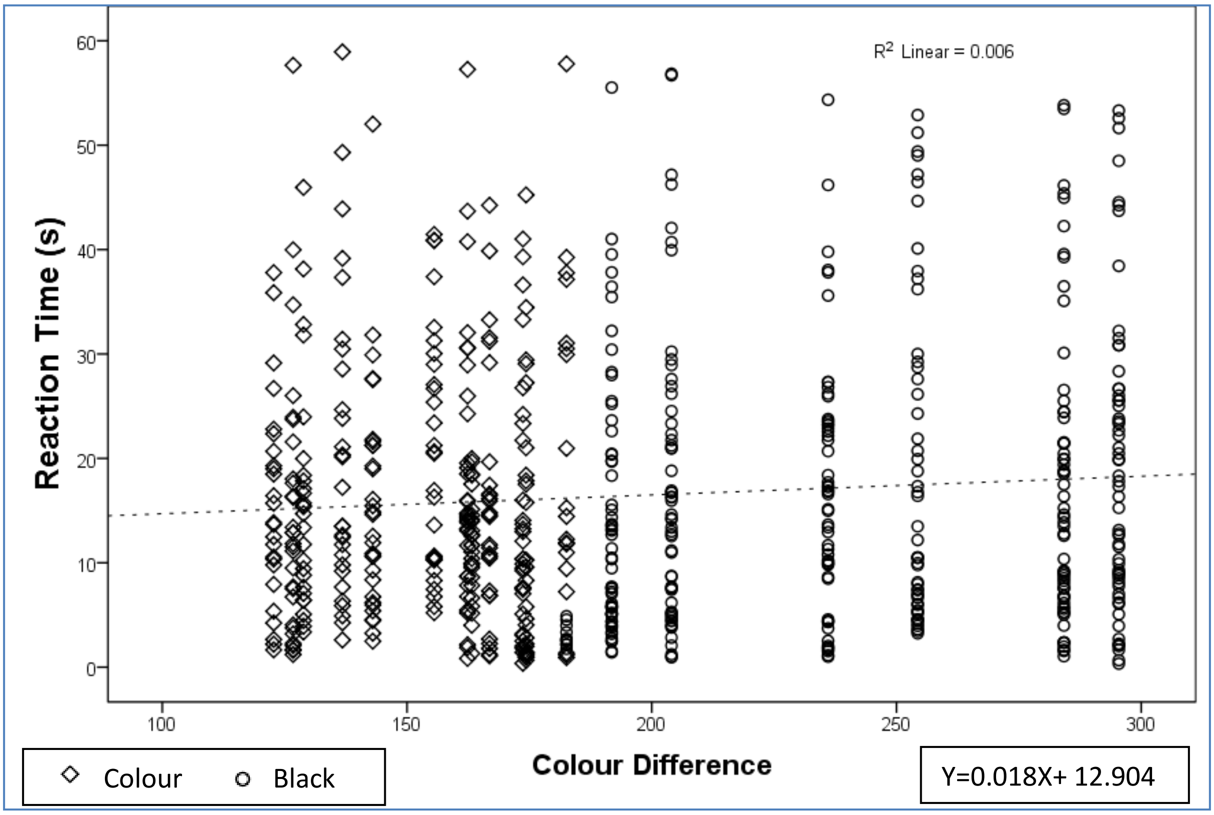
$$Y= 0.018\, X+ 12.904$$

$$R^2= 0.006$$

where:

- Y is the reaction time(s); and
- X is the colour difference between the foreground and background.

Additionally, the correlation between the colour difference and the number of fixations per second (fix/s) was studied and illustrated in Figure 8.5. The statistical analysis showed no correlation between the average fixation count and colour difference for the 24 map designs (P=0.209). Therefore, a regression model cannot be derived in accordance with the experimental conditions.



8.4: Scatterplots showing the influence of the colour difference on the users' performance and the related regression model

Because the complementary colour is the focal point of this research, colour legibility was studied in regard to the luminance of coloured labels. The colours (background and foreground) were chosen so that colour hue difference would be the highest when designing the map. However, as can be derived from Table 8.2, the luminance of these colours varies, which could also influence the readability of the colour combinations. As a consequence, the variations in luminance between the background and foreground colour will be evaluated and compared with the measured values: fixation duration.

The luminance difference between the foreground colour (label) and background colour was calculated from the measured Y-value in the XYZ-system:

$$\Delta Y = Y_{foreground} - Y_{background}$$

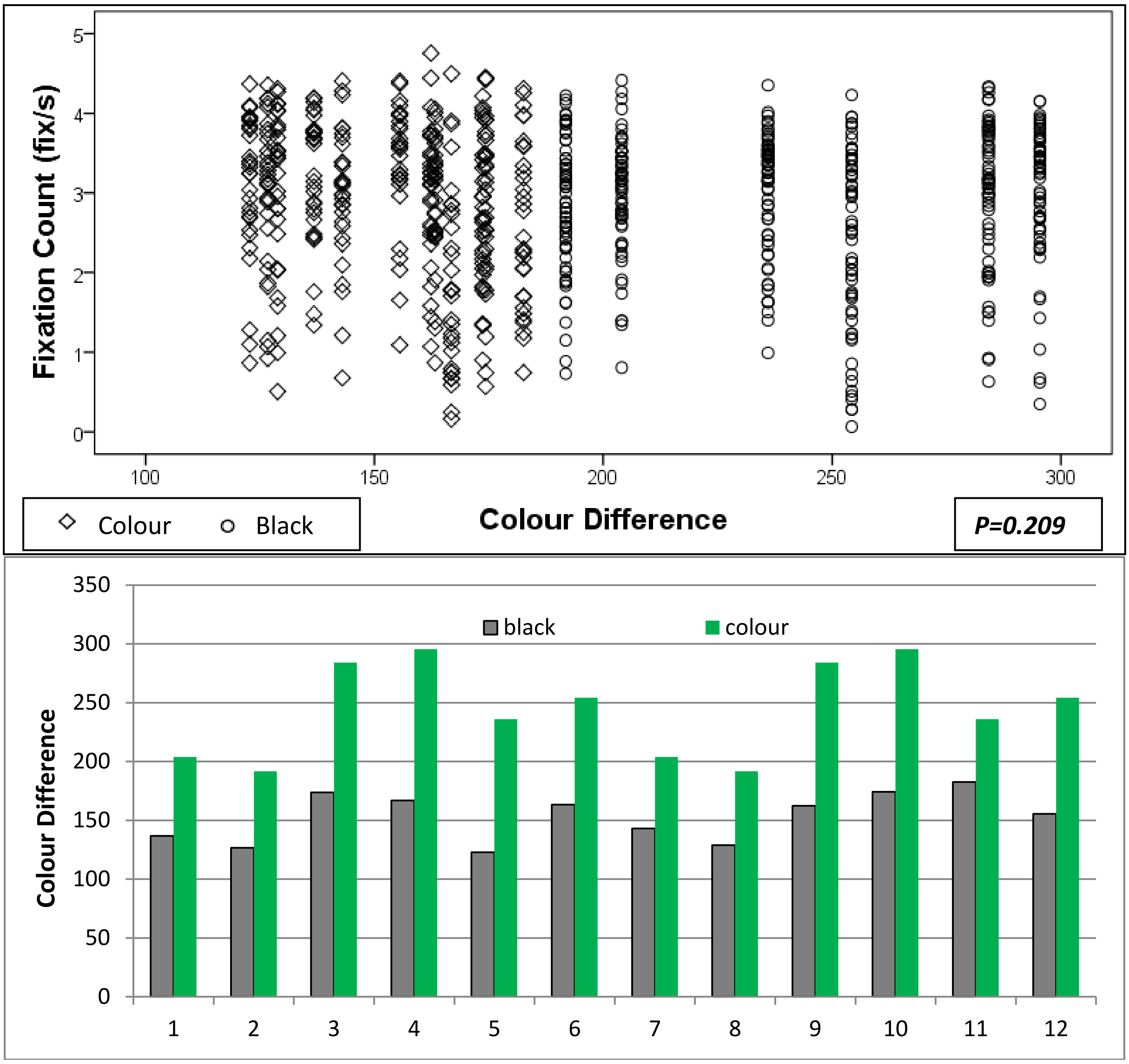


Figure8.5: Scatterplots showing the influence of the colour difference in relation to the average fixation count per second (up); colour difference between the foreground and background for both coloured and black labels (down)

The fixation durations in which the participants fixate on the target labels were extracted from the database and linked to the luminance difference between the target labels and their background. In Figure 8.6, a negative correlation was found between the luminance differences and the average fixation duration(s) ($P = 0.017$), i.e., more luminance contrast enhanced users' performance by reducing targets' fixation durations. It is worth mentioning that the order of the foreground/background colour is important when calculating the luminance differences.

The regression model is:

$$Y = 0.130\ X + 221.86$$

$$R^2 = 0.016$$

where:

- Y is the fixation duration(s); and
- X is the luminance difference between the foreground and background.

We have tested both the factual luminance difference, ΔY= Y foreground –Y background

and the absolute luminance difference ΔY= |Y foreground –Y background|. The absolute luminance difference showed no correlation with the fixation duration (P= 0.338).

To examine the influence of luminance difference on fixation duration, the analysis was performed only for the 12 coloured label designs. Since the black labels have nearly no luminance (Y foreground ≈ 0), the influence of black labels' luminance on luminance difference calculations is constant over the 12 background colour ΔY black labels= 0 –Y background (6), see Figure 8.6.

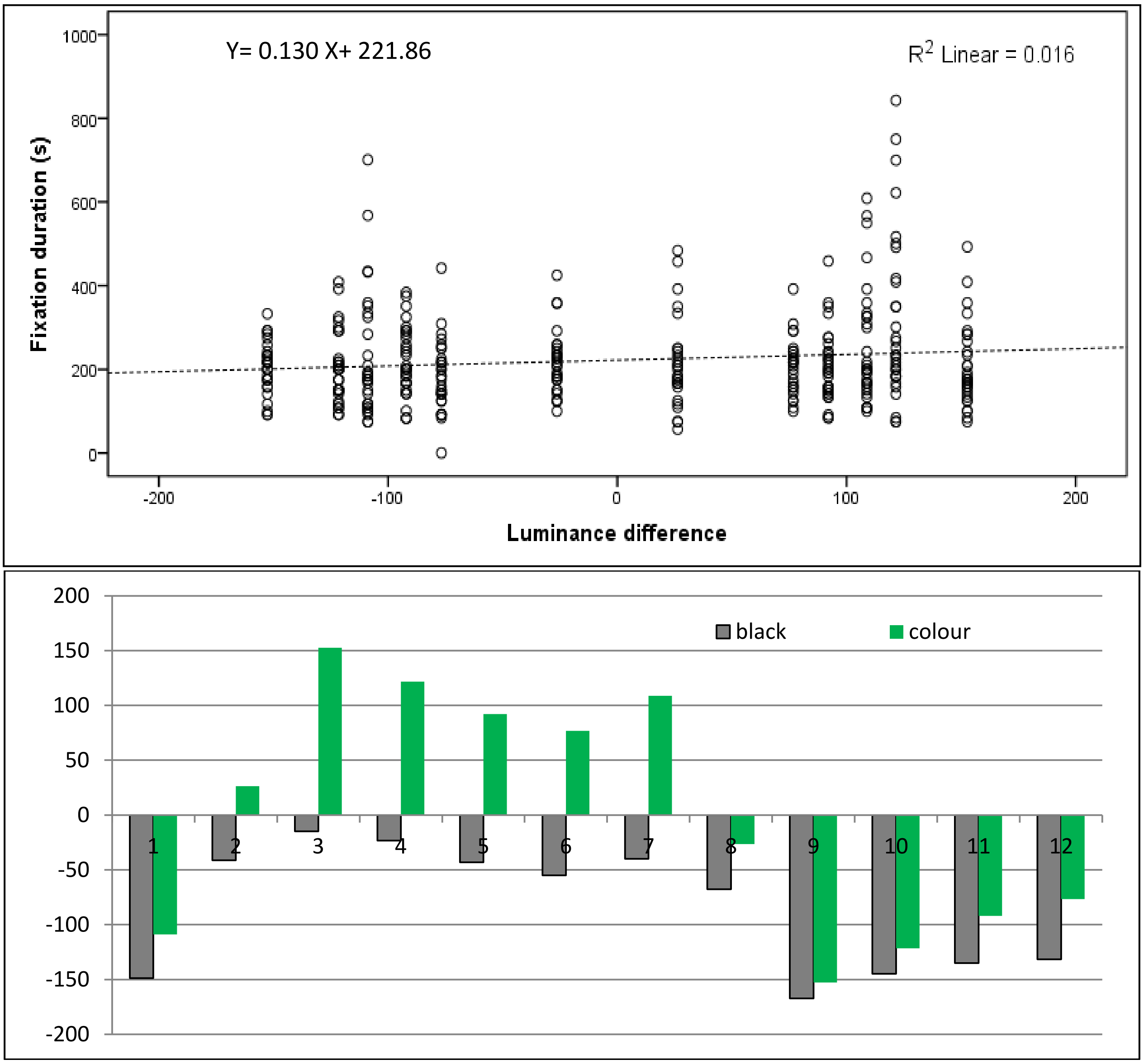


Figure8.6: Scatterplots showing the influence of the luminance difference in relation to the target fixation duration (up); the difference between the target luminance and background luminance for the 24 map designs (down)

8.3.5 Post-test questionnaire

After the search task, a questionnaire was presented to the participants to acquire their characteristics and preferences for both introduced designs (black versus colour). Out of the 15 experts and 16 novices, only one female expert and two male novices preferred coloured labels over black labels. They, in their open comments, explained that finding coloured labels was outstanding. The colours derived their search and helped in finding the targets. An ANOVA test was applied to compare the efficiency (search time) of the black label design versus the coloured label design. None of the participants showed a higher efficiency (less search time) for the coloured label design as the preferred design (F=0.004, P=0.949; F=0.000, P=0.991; and F=0.011, P=0.961, respectively). However, the rest of the participants' results (who chose the black label design over the coloured label design) did not show a significant difference in efficiency between the black label design versus and the coloured label design.

8.4 Discussion

This research presented an eye-tracking experiment with the aim of investigating the influence of the map labeling colour on the map users' efficiency. The interaction between the map components - and more specifically- the colours requires several design rules to ensure higher user efficiency. Within the user analysis, a constant trend was not shown between the black colour designs regarding user efficiency in the search task. The black labels' efficiency is not always less or more than the comparable coloured label. Only one significant difference showed that the user efficiency of reading colour12 (rose) on its complementary colour has a negative impact on the efficiency of reading the label. Meanwhile, the other design was not significantly different (rose in the background). These results are consistent with the findings of Garcia and Caldera (1996) and contribute to their suggestion of using different colours in the foreground for better legibility. However, it does not agree with Brewer *et al.* (1996) because labels with complementary colours that were used in the diverging schema do not prove to be more efficient than black labels.

The study of the reaction time measurements and fixation count per second shows that the number of fixations per second is independent of the users' overall efficiency, which in turn, refers to the fact that each stimulus influences its search rules. This result strongly agrees with Brandt & Strak (1997), Wolfe *et al.* (1989) and Strak & Ellis (1981), who stressed the importance of stimuli design on target search and interpretation. It is worth noting that comparisons (black -coloured) that included colour10 and colour12 showed the most frequently significant differences of the pairwise comparisons versus the studied eye-tracking metrics.

A between users analysis was tested through the MANOVA interactions. It showed no effect of the expertise and gender differences on the reaction time measurements, fixation duration and fixation count. This result confirms that black labels on the foreground have a similar cognitive load along with a different background in regard to the users' characteristics.

Regarding labels' designs, the results agree with the general theories related to colour differences and legibility (Williams, 1967; Shlaer, 1973; Johnson and Casson, 1995; Carter and Huertas, 2010; De Vries *et al* , 2013) because a correlation was found between the colour difference and reaction time measurements. However, no correlation was found between the colour difference and average fixation count. Having no correlation between the colour difference and fixation count is a twofold issue. First, the search on maps is a subjective issue (Phillips, 1981), and thus, a considerable variation between the number of fixations can be found between one participant and another. Second, the coloured objects are not abstracted or have no fixed geometrical shapes (like described in theories), but they are a design of functional words varying in shape over the 12 map designs because the size of the target labels was controlled by the number of their letter. Therefore, it cannot be concluded that the colour difference is irrelevant to legibility.

To examine the complementary colours, a fixation duration in which users fixate on the target label -when they located it- was extracted, and a statistical analysis was made. These factual fixation durations show a correlation between the target fixation duration and luminance difference. The results agree with Shlaer (1937). Our experimental conditions show a correlation between the fixation duration and luminance difference. Extensive map designs, which involve more colour contrasts, shall be studied in regard to the luminance differences and foreground/background colour differences. The order of colours (foreground versus background) proved to be very vital considering label legibility. Therefore, it is important to take into account the colour in the foreground in regard to the background colours when designing maps (e.g., the legibility of yellow foreground - blue background does not resample the legibility of blue foreground - yellow background).

The experiment showed that a minor ratio of users preferred coloured labels (approximately 10%) over black labels. However, the users' preference (of either design) was not correlated to their performance because their efficiency was not affected by their selection for the coloured or black design.

8.5 Conclusion and future work

The label colour was analyzed in combination with its background colour. The complementary coloured label and the corresponding black label were compared pairwise over the 12 designs. In addition, two users' groups were distinguished based on gender and experience. The results demonstrated several differences on the colour combinations concerning user efficiency while performing the search task and no significant differences regarding their characteristics. More luminance contrast and colour difference enhanced users' efficiency by reducing fixation duration and reaction time. Meanwhile, the preference in colour design did not cause higher efficiency. The research results are limited to the experimental condition, especially the screen type and its calibration. As part of a large study, colour hue was described in this chapter, but the study did not cover all aspects of colours. Thorough experiments will be conducted to demonstrate the differences in colour value and saturation where small and large colour differences will be compared. These results can be employed in screen design and more specifically in map label design to provide higher user reading efficiency. The research was based on a basic map design and will be extended to cover topographic maps where the use of colour is functional and standardized over the world. Moreover, different screen sizes need to be studied as well.

8.6 References

Anderson, G. M., Hienke, D., & Humphreys G. W. 2013. Top-down guidance of eye movement in conjunction search. Vision Research, 79, pp.36-46.

Arditi, A. & Cho, J. 2007. Letter case and text legibility in normal and low vision, Vision Research, 47 (19), pp.2499-2505.

Arditi, A. & Cho, J. 2005. Serif and font legibility, Vision Research, 45 (23), pp. 2926-2933.

Benjamins, J. S., Hooge, I.T.C., van Elst, J. C., Wertheim, A. H., & Verstraten, F. A.J. 2009. Search time critically depends on irrelevant subset size in visual search. Vision Research, 49 (3), pp.398- 406.

Bernard, M. L., Chaparro, B.S, Mills, M. M., & Halcomb, C. G. 2003. Comparing the effect of text size and format on the readability of computer-displayed Times New Roman and Arial text. International Journal of Human-Computer Studies, 59 (16), pp.823-835.

Bertin, J. 1967. Semiology of Graphics: Diagram, Networks, Maps, Translated by William J. Berg, 2011, Esri press, California.

Brandt, S. A. & Strak, L. W. 1997. Spontaneous Eye Movements During Visual Imagery Reflect the Content of Visual Scene. Journal of Cognitive Neuroscience, 9(1), pp.27-38.

Brewer, C. A. 1996.Prediction of simultaneous contrast between map colors with Hunt's model of colour appearance, Colour Research and Application, 21(3), pp.221-235.

Brewer, C. A., MacEachren, A. M., Pickle, L. W., & Herrmann, D. 1997. Mapping Mortality: Evaluating Colour Schemes for Choropleth Maps, Annals of the Association of American Geographer, 87(3), pp. 411-438.

Carter, R., & Huertas, R. 2010. Ultra-Large Colour Difference and Small Subtense. COLOUR research and application, 35(1), pp.4-17.

Chan, A. H. S. & Lee, P. S. K. 2005. Effect of display factor on Chinese reading time, comprehension score and preference, Behaviour & information technology, 24(2), pp. 81-91.

De Vries, J. P., Hooge, I.T.C., Wertheim, A. H., & Verstraten, F. A. J. 2013. Background, an important factor in visual search. Vision Research, 86, pp.128-138.

Deeb, R., Ooms, K. & De Maeyer, P. Typography in the Eyes of Bertin, Gender and Expertise Variation. The Cartographic Journal, 49(2), pp.176-185.

Duchowski, A. T. 2007. Eye Tracking Methodology, Theory and Practice. (2nd ed.). Springer, London.

Foley, J. M., Varadharajan, S., Koh, C. C., Farias, M. C. Q. 2007. Detection of Gabor patterns of different sizes, shapes, phases and eccentricities. Vision Research, 47 (1), pp.85-107.

Garcia, M. L. & Caldera, C. I. 1996.The effect of colour and typeface on the readability of on-line text. Computers and Industrial Engineering, 31(1), pp.519-524.

Haklay, M. M. & Nivala, A. M. 2010. 'User-centred Design', In: Interacting with Geospatial Technologies. ed. by Haklay, M. M. chapter 5, pp.91-106, John Wiley& Sons. Chichester.

Harrower, M. 2007. The cognitive limits of animated maps. Cartographica, 42(4), pp.349-357.

Hegarty, M., Canham, M., & Fabrikant, S. I. 2010. Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. Journal of Experimental Psychology: Learning, Memory, and Cognition, 36(1), pp. 37-53.

Huang, D-L, Rau, P-L. P, & Liu, Y. 2009. Effects of font size, display resolution and task type on reading Chinese fonts from mobile devices. International Journal of Industrial Ergonomics, 39(1), pp.81-89

Jacob, R. & Karn, K. 2003. 'Eye tracking in human-computer interaction and usability research: Ready to deliver the promises'. In R.Radach, J.Hyona & H.Deubel (eds.), The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research, pp.573-605, Amsterdam: Elsevier.

Johnson, C. A. & Casson, E. J.1995. Effects of luminance, contrast, and blur on visual acuity, Optometry and Vision Science, 72(12), pp.864-869.

Nachmias, J. 2011. Shape and size discrimination compared. Vision Research, 51 (4), pp.400- 407.

Nielsen, J. 1993 Usability Engineering, Academia Press, San Francisco.

Olson, J. M & Brewer, C. A. 1997. An Evaluation of ColourSelection to Accommodate Map Users with Color- Vision Impairments, Annals of the Association of American Gaeographer, 87(1), pp.103-134.

Ooms, K., De Maeyer, P., Fack, V., Van Assche, E. & Witlox, F. 2012. Interpreting Maps Through the Eyes of Expert and Novice Users, International Journal of Geographical Information Science, 26(10), pp.1773-1788.

Phillips, R.J. 1981. Search for Target in a Random Arrangement of Names: an Eye Fixation Analysis. Canadian Journal of Psychology, 35(4), pp.330-346.

Rosenholtz, R., Nagy, A. L., & Bell, N. R. 2004. The effect of background colour on asymmetries in colour search, Journal of Vision, 4, pp. 224-240.

Rubin, J. & Chisnell, D. 2008. Hand Book of Usability Testing: How to Plan, Design, and Conduct Effective Tests, 2nd edition, Wiley, Indianapolis.

Shlaer, S. 1937. The Relation Between Visual Acuity and Illumination, Journal of General Physiology, 21(2), pp.165-188

Slocum, T. R, McMaster, R. B, Kessler, E. C, & Haward, H. H. 2005. Thematic Cartography and Geographic Visualization, 2nd edition, Person Prentic Hall, Upper Saddle River, NJ.

Strak, L. W. & Ellis. S. R. 1981. 'Scanpaths revisited: cognitive models direct active looking'. In Fisher, D.F *et al.* (eds.), Eye movements: Cognition and visual perception, Lawrence Erlbaum Associates, Hillsdale, NJ, pp.193-226.

Werman, M. 2012. Improving Perceptual Colour Difference using Basic ColourTerms. Computer Research Repository, pp.1–14.

Williams, L. G. A. 1967. Study of visual search using eye movement recordings. Honeywell Document 12009-IR 3, Navy Contract NONR 4774(00).

Wolfe, J. M. 1994. Guided search 2.0: A revised model of visual search. Psychonomic Bulletin & Review, 1(2), pp.202-238.

Wolfe, J. M., Cave, K.R. & Franzel, S. L.1989. Guided search: an alternative to the future integration model for visual search. Journal of Experimental Psychology: Human Perception & Performance, 15, pp.419-433.

9

General Discussion

This chapter presents an overview of the research objectives and review the main answers to the research questions (Section 9.1). Then, a discussion of the combination between the objective and subjective cartographic text design will be presented (Section 9.2). Finally, the applications and challenges of cartographic text, which cartographers face, are considered in Section 9.3.

9.1 Summary

The symbology of the cartographic text issued some considerations in regard to map design, map use, map production and display, and map users. Attempting to convey the implementations of cartographic text design and explore their nature, four research objectives were assigned:

- **Research Objective A:**
  Improve both the esthetical and the functional design of cartographic text by the employment of the visual variables on screen map design.
- **Research Objective B:**
  Investigate the influence of cartographic expertise and gender differences on the map users' preferences and performance towards cartographic text designs.
- **Research Objective C:**
  Discuss the relationship between cartographic text colour and map background colour design.
- **Research Objective D**
  Investigate the impact of lettering systems on cartographic text design.

The four research objectives were rephrased and addressed in five research questions which were reflected in the dissertation's chapters. Each chapter attempts to answer two or more of the research question as defined in Table 9.1. The table indicates the link between the research objectives and the research question and how they were presented in each chapter of this dissertation. Additionally, the outline of the research design is presented regarding map design, data collection methodology and technique, the tests' design (within and/or between users' design), and the main results concluded in each chapter.

The research answered empirically to the research question as follows:

Table 9.1 : The head line of the dissertation's component

Ch	RO	RQ	Map design	Testing technique	Participants	Highlights
Ch2	RO: A RO: B	RQ 1 RQ 2	Thematic & topographic Blank background	Questionnaire	Novices vs. Experts Female vs. Male	* Visual variables can be functionally applied on cartographic text * The complexity of the design has an influence on users' preference * The preference of individual variables has an excellent agreement between users' group
Ch3	RO: A RO: C	RQ 1 RQ 4	Thematic & topographic Blank, grey scale, hot colour, cold colour background	Questionnaire	Novices & Experts Female & Male	* Maps' background colour has no influence on typographic preference * Excellent agreement on the use of size, shape, orientation, and texture * No individual design difference
Ch4	RQ 1 RQ 2	RQ 1 RQ 2	Thematic &topographic Blank background for thematic and colour for topographic	Questionnaire Time measurement	Novices vs. Experts Female vs. Male	* Label design enhance map readers' efficiency * Few significant differences between users 'group efficiency wise * A positive correlation between the design complexity and users' efficiency
Ch5	RO: A RO: D	RQ 1 RQ 3	Thematic & topographic Blank background for thematic and colour for topographic	Questionnaire	Novices & Experts Female & Male	* Latin and Cyrillic cartographic text did not show any significant difference preference wise * Users' preference of cartographic text design regarding Latin, Arabic, and Chinese did show significant differences between groups
Ch6	RO: A RO: D	RQ 1 RQ 3	Thematic & topographic Blank background for thematic and colour for topographic	Questionnaire Time measurement	Novices & Experts Female & Male	*Together, Latin, Arabic, and Chinese cartographic text did not show a trend on RTM * Between-users' designs: cultural differences influenced users' efficiency * Within design: some visual variables influenced significantly users' efficiency
Ch7	RO: A RO: B RO: C	RQ 1 RQ 2 RQ 5	Thematic Coloured background	Questionnaire Eye tracking	Novices vs. Experts Female vs. Male	*Tracking users' search strategies * Employment of multi-package open source software * No definition of search strategies was obtained over the stimuli
Ch8	RO: A RO: B RO: C	RQ 1 RQ 2 RQ 4	Thematic Coloured background	Questionnaire Time measurement Eye tracking	Novices vs. Experts Female vs. Male	* A negative correlation between labels' legibility and their luminance difference with background colour * Black labels is not more efficient than coloured labels in relation to the background colour * Users' preference of colour design does not affect their efficiency

RQ1 *How can the visual variables be applied to cartographic text, and more specifically map labels?*

Visual variables can be used to guide a 'good' label design by applying the appropriate visual variable which suit labeling functions.

In his dissertation 'The Look of Maps: An Examination of Cartographic Design', Robinson mentioned that the application of lettering is one of the most difficult issues in cartography (Robinson, 1952). He thought of map lettering as external map element. Some cartographers agree with him regarding his map definition, which explicitly exclude letters because they are not visible on a conventionalized depiction of the earth. Other cartographers strongly disagree and consider lettering as a symbol type on the map in addition to point, line, and area (Fairbairn,1993; Tyner, 2010). Based on this consideration, the visualization of cartographic text can follow the same rules of the visualization of the other symbol types.

Watzman (2002) defined typography as the art of defining and arranging the general appearance of type. Additionally, the evolution of cartographic text is an interesting study of the interaction of art, tradition, and convention Robinson (1952). In this dissertation, both the application of visual variables and the consideration of Gestalt lows were implemented to set the design of cartographic text. Chapter 2 explained how visual variables can be applied on the typographic elements, the function of this application, the perception properties and the aesthetics values. Our usage of size is based on our definition and its linkage to our maps' scale (1:10 000), therefore, we suggest that our results are applicable in the case which preserve the map scale and point size or the proportion between them.

Although the sharp angles, firm edges, and complex curves of letters' forms make them one of the most complicated visual elements, much of this complexity hides unnoticed because the shapes of letters are well known to the readers (Robinson, 1952). Cartographers acknowledged that the cartographic text shape is infinitive (Robinson, 1952; Bartz, 1970; MacEachren, 1995; Tyner, 2010). Yet, dividing the shapes of cartographic text into classes or categories, considering the fonts' essential characteristics, facilitates map depiction and uses the 'optimal' shapes for different functions. Different shapes of cartographic text have the characteristics that depict features which may either harmonize well or contrast well with the characteristic of feature shapes on a map.

Notwithstanding, in cartography the choice between upright (straight) and italic is usually based solely on the desire to differentiate between categories of data by change in type face. Yet, italic type is generally undesirable except for emphasis (Robinson 1952). The research is in coherence with Robinsons' description of italic cartographic text. Although he described it for paper mapping, the results agree that users are in favour of using straight cartographic text more than using italic text, even when using screen maps.

In accordance with the most common orientation (north up) of maps our results showed positive responses towards the horizontal orientation of labels. This finding contributes to Robinsons' observations (1952) that when the letters are presented on a non-horizontal base they will attract attention by their inharmonious position and appearance. Although non-horizontal positioning of letters attracts attention, so many non-horizontal labels distract attention and this issue could have caused a delay of response time measurement in relation to horizontally placed labels. The study showed that users are in favour of horizontal orientation more than other sort of label placements (tilted orientation, i.e., along with the diagonal of the polygons' shape or mixed orientation, i.e., combines between horizontal and tilted).

The texture of cartographic text (labels) can be built by the combination of two or more visual variables. Therefore, when designed functionally, the cartographic texts help to correctly perceive the attributed features. The complexity of cartographic text design is most functional with texture design. The tested designs of labels' textures showed less preference and less efficiency with more complex labels' texture. Research on symbolization design complexity agrees with our finding considering memory of general pattern, and communicating general patterns (MacEachren, 1982).

In contrary to Bartz (1970), who suggested using the typographic literature, which studied the typographic design on plain test, for cartographic text design to improve maps' typography, this research provided empirical confirmations that cartographic text has their own rules of design to provide efficient and preferred cartographic text design. Yet, some typographic design principles on plain text are still applicable on cartographic text such as the findings of Shieh *et al.* (1997). They demonstrated how type face significantly affected visual performance, as participants performed better with computer type than with the aesthetically more pleasing Kia type.

The fact that there will be differences in legibility and perceptibility between various letter combinations is evident (Robinson 1952). The employment of visual variables on the cartographic text were studied in the dissertation's chapters (chapter 3-8) to achieve good typography, as good cartographic text is the base of good visual design infrastructure (Watzman, 2002).

Two aspects of usability were studied on this dissertation. First, users' preference towards map typography was presented in chapters 2, 3, and 5. Preference responses consist of the subjective data that measures participants' opinion (Rubin and Chisnell, 2008). Noticeably, users' preference showed patterns when size, shape, orientation, and texture were applied on labels when Latin lettering system was studied in relation to different background colours, but when different lettering system were compared, the studied parameters of visual variables significantly influenced users' preference.

Chapters 4, 6, 7, 8 discussed users' performance towards these designs through locating the assigned target labels. Performance data, which was collected through users' efficiency, was the objective measures of their behaviour (Rubin and Chisnell, 2008).Chapter 8 indicated also the effectiveness of colour designs through fixation duration and fixation count analysis.

RQ2 ***How do user characteristics influence the perception of label design?***

Users' expertise and gender differences can significantly influence the perception and usability of label design.

The individual's performance, in cartographic language can be affected by his own characteristics including intelligence, visual perception, map-reading ability, drawing ability, home environment, previous experiences with maps and attitudes toward maps (Gerber, 1981). Map users' characteristics steer somehow the overall responses to special tasks (Nielsen, 1993; Dumas and Redish, 1999; Rubin and Chisnell, 2008, and others), therefore two user test designs (Both between and within users' design) were considered to verify the effect of users' characteristics.

To address this research question, we specifically focused in chapter 2 on users' characteristic influence on label preference, when two users' groups gender and expertise. The variations were distinguished for the between users' design where females were tested against males, and novices were tested against experts.

When studying labels' variations in size, shape, orientation and texture, few significant differences were located between females and males considering users' preference of the application of texture variable. This difference was located for the most complex texture design. This research question was also tackled in chapter 4 when we studied the influence of users' characteristic onto their efficiency regarding label design. The results are compatible with the result of chapter 2 as few significant differences were located. Noticeably, in comparison between gender and expertise groups, less significant differences in the expertise groups appeared considering both users' preference and efficiency. But when the variable colour hue was studied (chapter 8) no influence of users' group on the results was shown between both gender and expertise groups. Since we used a different methodology for each chapter, i.e., questionnaire for chapter 2, online reaction time measurement in chapter 4, and reaction time measurement through an eye tracking experiment, these results highlight on the fact that the stimuli itself had a significant effect on the results. Our previous statement is also supported by the fact that our results contradicted many studies that were able to locate significant differences between users' groups (Hambrick and Engle, 2002; Ooms *et al.,* 2012; Ooms *et al.,* 2013; and Popelka & Brychtova, 2013). Chapter 8 showed no significant difference between users' group while eye tracking data was analysed, but chapter 2, 4 showed few significant differences between users' groups. The reason for such results could be related to the analyses methodology as we used one-sample T test in chapter 2 and ANOVA and one-sample T test in chapter 4.But in chapter 8 we used MANOVA where the test has larger statistical power. Therefore, we may say that the statistical differences which were located in chapter 2 and 4 could have occurred by chance.

The research moved from the quantitative analysis to the qualitatively analyses (chapter 7) by studying scanpaths. We used sequence analysis methodology to investigate search patterns between groups, which showed no determent patterns between both gender groups and expertise groups. Since few patterns appeared when scanpaths were analysed, we would think of the complexity of the task as a negative aspect that prevented using this technique to study pattern differences between groups. Although the results could not explain why and how these few differences in the users' groups occurred, especially that no consistent patterns appeared over the analysis (both quantitative and qualitative), we do agree that users' characteristic should be considered (Nielsen, 1993; Dumas and Redish, 1999; Rubin and Chisnell, 2008, and others)when studying the usability of map designs.

RQ3 ***How do lettering systems affect the usability of visual variables?***

The embedded characteristics of lettering systems signifficantly influence label design and how we can use visual variables for both preferred and efficient design.

The alphabet is the best tool that human kind has for storing thoughts, ideas, and instructions until they can be employed. Yet, images are preferred to wards because well designed images are direct, well understood, and can avoid the complications of language (Shirreffs, 1992). Maps are a homogeneous composition of both graphs and alphabets and thus well designed maps are good communicational tools that can communicate complex messages using graphic techniques, which overcome the barriers of verbal language.

Since map textual elements hold essential functions (Fairbairn, 1993), map image cannot be completed without the cartographic text, a critical communication issue of map language appears. What lettering system is used to label and explains features over the map and how they are used on the map face is a vital factor to consider when designing maps. To clarify this issue, chapter 5 discussed users' preference of four of the most used lettering systems (Latin, Cyrillic, Arabic, and Chinese), and chapter 6 discussed efficiency wise the application of cartographic text of three diverse lettering systems (Latin, Arabic, and Chinese). Designers and typographers have long battled about the appropriateness of certain design mannerisms. In both chapter 5 and chapter 6, the characteristics of Latin, Cyrillic, Arabic, and Chinese lettering systems were implemented to optimize the use of visual variables for cartographic text design.

Although cartographers argued considering the cartographic text as a symbol (e.g., Robinson,1952 ; Fairbairn, 1993), we addressed the four distinguishable lettering systems as symbols and we parallely applied Bertin's variables on these symbols. When the designs of Latin and Cyrillic were compared preference wise in chapter 5, no significant differences were located and this could have resulted essentially because both lettering systems encompass similar characteristics. Therefore, the research continued to compare preference wise between three different lettering systems, which have different characteristics (Latin, Arabic, and Chinese), where significant differences were located regarding cartographic text design. Additionally, the research in chapter 6 was able to locate significant differences regarding the efficiency of users, when these lettering systems were tested.

Regarding size, shape, orientation and texture of cartographic text design, between and within lettering systems analysis showed significant differences in users' responses preference and efficiency wise. These results can guide map designer to design both monolingual and multilingual maps. When map designers aim to equalize users' efficiency they can use the variables that ware coherent with labelling goals and which can achieve similar efficiency (reflected in reaction time measurement).

Two major characteristics in Latin were discussed and compared to similar characteristics in Arabic and Chinese. Serif and sans serif in Latin were compared to similar fonts in Arabic and Chinese which resample serif and sans serif fonts in view point of Arntson's (2011) explanations of these characteristics, where 'serifs are the finishing strokes in all letters other than O and Q. Serifs are either unilateral, protruding in one direction only, like the top left serif on F; or they are bilateral, as when they protrude in two directions, like the bottom serifs on F. And the sans serifs are stripped to the bare minimum by losing the serif appendages'. Although it is not common to speak of serif and sans serif fonts in both Arabic and Chinese lettering systems, the previous definition demonstrated some similar characteristics of their typographies considering letter (finishing) strokes' design. Our experimental results do prove that these lettering systems influence and guide the use of visual variable in typographic design efficiency and preference wise.

Although the experiments' conditions did not test the three lettering system on one map, which could have made our results an empirical reference for multilingual mapping, we suggest the use of our finding as guide to define users' efficiency of each lettering system individually and combined, till the otherwise is proven.

RQ4 *How do maps' background designs influence user preferences and performance together with the label design?*

The colours' hue and colours' luminance of maps' background can influence the usability of labels design including labels' colour.

Today, with the expanding use of screen maps and the considerably cheap printing coasts, coloured map have become ubiquitous. Colours allow great flexibility in map design as they facilitate the distinction between figure –ground elements (Tyner, 2010). Although the use of colour is considered desirable, we must be aware that its implications are challenging (e.g., later inhabitation explained by MacEachren, 1995). Colour can be used to enhance the effectiveness of graphical display (Christ, 1975; Murch, 1985). This enhancement is guided by users' preference as preference is a factor that aid comprehension (Kaplan, 1987). Colours are more preferred over black and white designs (Brewer, 1996), and while presenting beauty, colour is one of the strongest components in the preference space (Schenkman and Jonsson, 2000). Yet, the use of colour in visual display should not only be influenced by the subjective preference of the designer or users, but also by the cognitive perceptual constrains of the users. (Ling and van Schaik, 2002). Because when colours are used properly, they can lead to faster search times (Christ, 1975).

Chapter 8 combined two of the most criticized aspect of map design, colours in addition to map typography (Tyner, 2010). It attempted to attribute both users' preference and performance of figure – ground interaction, since coloured cartographic texts were presented as figures and the diverse background colours were presented as the ground. And because the transition from 'performance-related' to more 'pleasure-fulfilling' activity is essential to evaluate what pleases and attract map users (Schenkman and Jonsson, 2000; Nielsen, 1993). Therefore, the chapter presented further link between users' preference of cartographic text colour and users' efficiency of these colours were made. No correlation between users' preference and users' efficiency was found since users' preference of coloured labels was not linked to better efficiency in comparison to black labels' efficiency.

Colour is presented in several systems which quantify colours (e.g., RGB, CYMK, HSV, HSB, CIE, Munsell). These different quantifications illuminate on different aspects of defining and analysing colours. The idea of using a certain-dimensional system could help to identify colour but colour analyses should go beyond one colour definition in certain systems. The conversion between systems is a useful tool to explore different colour dimensions and their influence on the analyses.

Although complementary colours proved to be more efficient in comparison to other colour comparison (Brewer, 1996), our results showed that colours' luminance differences is more powerful than colours' complementary designs. A negative correlation between users' reaction time and luminance difference was found, i.e. higher users' efficiency corresponded to more luminance difference. Chapter 8 support the theory in which the search discriminability of targets is related to colour difference between targets and background items and legibility is related to luminance contrast between the letters and their adjoining background (Hecht, 1928; Shlaer, 1937; Carter and Huertas, 2010; Carter and Silverstein, 2010).

An important factor of studying screen maps is monitor calibrations. Normally map designer identifies his pallet based on the numerical definitions presented by the design software.

Unfortunately, map designer cannot guarantee that all map users will see the similar colour that he initially defined, because what colour users see on their screens is redefined by the screen calibrations themselves.

Our observation of users' preference regarding the colours of the map on the background when using black labels on the foreground demonstrated that users' preference of label designs will not be influenced by the colour in the background, when labels were depicted with black. Additionally, when coloured label were compared to black labels, users preferred black. Although this result could have been affected by the presented stimuli design (complementary colour stimuli), it also can be explained by the fact that black colour use is the conventional and traditional use for labelling.

RQ5 ***What type of search strategy do users follow when they search for labels over the map face and for certain map design conditions?***

Map users' search strategies are considerably inconsistent when we consider the same users' search strategy and multiple users' search strategies for different stimuli.

One of the earliest attempts to understand map users' search strategies was thoroughly investigated by Phillips *et al.* (1978) and followed by Phillips (1981) when he used the eye tracking technique to study the cartographic text reading and its controlling and influencing factors. Their collection of target labels included variations in letters assembly to write the labels, the neighbouring labels to the target labels, labels size, label position on the map and the complexity of maps. This early attempt to understand the search strategies on paper maps could not identify patterns or basic manners which can describe the search strategies. Philips' study was followed later by many which used eye tracking technique to understand the way that user search on screen maps to locate targets (Wolfe, 1994; Ooms *et al.,* 2012; Ooms *et al.,* 2013; Popelka and Brychtova, 2013; and others). The similar technique was used to investigate the search strategies on screen land use map, where the research incorporated the variables that Phillips *et al.* (1978) and Phillips (1981) used in addition to labels design. To analyse the eye tracking qualitative data (scanpaths), multi-packages open source software were used in chapter 7, which identified the similarities and the dissimilarities between scanpaths.

Although the analysis tried to quantify the qualitative eye tracking data to find search patterns, our stimuli seemed to be too complicated to show distinguishable search pattern over each stimulus and the overall stimuli. A sum of 114 labels was presented in the stimuli and it was observed during the testing procedure that users', sometimes, scan the map twice or thrice or even more when the participant was distracted. This issue created extra scanpaths, e.g., from A to E several times, which could have conflicted the analysis made by OGAMA. In spite of all the drawback of the stimuli design, the analysis was capable of generating a model that describes the variations between users' search strategies.

Additionally, an important factor of the presented methodology is how to identify areas of interests (AOIs), which can be either generated automatically or defined by the researcher. When generated automatically, squares of AOIs will be defined, but the manual creation will follow the subjective/objective thoughts of the researcher. Although we complicated the analysis, to some extent, by creating 16 area of interest, we trust that it is better choice to go for manual creation of AOIs because they will be compatible with the basic features of the map.

Because it is vital to know how map users' scan maps in their search for targets (labels), which can guide map designer to better design, many more techniques can be used to analyse scanpaths data (Andrienko *et al.* 2009, 2012). The prospective research, will consider map type and how the basic map elements are depicted (e.g., a long linear feature could direct users' attention).

9.2 Improving cartographic text design on screen maps

The map is both an objective and subjective products. The decision of what to include on a map is related to many objective factors, represented in the map purposes themselves, but also to subjective factors, such as what the cartographer, the mapping agency, or the client want to show (Tyner, 2010). Additionally, to evaluate map products, both subjective and objective aspects are considered.

While both objective and subjective measurements of users' response to cartographic text designs were the focus of this dissertation, an integration of both measurements was not discussed so far.

On the one hand, the objective measurement is presented by users' efficiency regarding certain design. The subjective measurement is presented by users' preference and users' characteristics differences, on the other hand. The 'optimal design' is the combination of objective and subjective cartographic text.

To address the objectivity of the cartographic text design, time measurement technique was used to define the efficiency of each application of the visual variable (including size, shape, orientation, and texture). Additionally, users' preference, as the subjective scale, was used to indicate the beauty of screen map as it is an important factor determining how it will be experienced and judged (Schenkman and Jonsson, 2000). Along the presented usability experiments, the combination between both measurements as not possible for size, shape, texture, and colour. The reason for this is that the most efficient variable was not the most preferable variable. But for the overall orientation of labels, horizontal orientation was the optimal design as it was the most efficient and preferable orientation parameter.

Although the experiments were not able to define an 'optimal' design for size, shape, and texture, they defined the subjective and objective parameters that can help map design in regard of producing either attracting maps or more functional maps. Yet, a comparison between users' efficiency was presented, where users' efficiency of the studied variables are addresses. Additionally, users' preference regarding the used parameter is defined, which can assist designers for their choice of both the objective and the subjective parameters. These findings correspond to Lliinsky (2010), who found that the graphical efficiency is a necessary, but not solely sufficient, ingredient to achieve beauty. Our study proved that vies versa is also true regarding cartographic text design. Therefore we can say that text beauty does not always comply with text efficiency.

When considering users' characteristics, few significant different regarding the between- group analysis was located, therefore, the subjective differences of the cartographic text design can be neglected at this level of measurement.

Considering the interaction between map's background and foreground colour, an important improvement of cartographic text design is presented in chapter 8. It illuminates on the fact that colour difference enhances users' efficiency and luminance contrast can also enhance users' efficiency. However, users' preference of the cartographic text design is not influenced by the design of background colour. Therefore, to achieve the 'optimal' cartographic design, considering background-foreground colour interactions, cartographer shall focus on the efficiency of the colour design.

9.3 Limitation, challenges, critical reflections and avenues for future research

The studied topic is somewhat virgin regarding empirical proofs of the 'optimal' cartographic text designs on screen. Except what Bartz (1970) and Philipps (1978, 1981)presented for cartographic text for paper maps, what was presented before is a group of none proved suggestions considering mapping (Robinson, 1952; MacEachren, 1995; Slocum *et al.,* 2005; Kraak and Ormeling, 2010; andothers). These suggestions were based solely on cartographers' practice and observations. The topic's field was open for choices, but for each step of the research we had to make decisions, that define and limit our research and leave the door open wide for prospective researches.

Within the research, the investigation of the visual variables application was made in two steps. At the first step, cartographic text size, shape, orientation and texture were considered. At the second step, cartographic text colour-hue- was studied. Noticeably, Bertin's variable (1967) 'value' was not tested. The reason for excluding this variable from the series of experiments, presented in this research, is that the variable is critical in nature and does not comply with Withycombe's (1929) essentials of typography. It is not common to use light cartographic text on mapface, yet, it worth investigating some typographic conditions that involve colour value, such as the extreme difference of value between colours in the background and label colour in the foreground.

The dissertation presented a series of experimental studies that evolved from studying blank map to studying topographic colourful maps. All the experiment used screen maps that were presented to users at a fixed screen size. An important aspect of digital map size was not discussed thoroughly in the dissertation which is map resolution. The test used thematic maps whose size is 150kb and topographic maps whose size is 350kb. The use of these sizes guaranteed similar loading time for the online experiment. It shall be mentioned that the image size (maps in our case) affect the component clarity and legibility (as well as the resolution of the screen). For this reason, map component size can be measured by the number of pixels they have. Such use can give another definition for letter size. On the same map, cartographic text can be measured by the number of pixels they have instead of their measurement in points. This kind of measurements contributes to the map legibility issues considering map rasterization.

It is not clear for us whether we can apply the research findings on paper map. Such condition was not tested and the comparison of users' responses to the cartographic text designs on paper maps versus screen maps was not made before. Not only the presented material (paper versus screen), but also the viewing distance can control map usage conditions. The later issue conflicts the definition of cartographic text size and their related texture.

Within the boundaries of this dissertation, cartographic text design was studied by investigating application of Bertin's variables on labels and users' responses towards these designs. We initially addressed that we considered label as representative of all the categories of map textual element (Title, legend, supported information, etc.). Yet, we would invite map designer to use our generalization when designing their maps but also to consider the characteristic of map title, map legend and the rest of textual element on the map.

As our studied maps are two dimensional (2D), we think that some of our results could be applied safely on three dimensional (3D) static screen maps. But, because of the nature of the three dimensional maps, the cartographic text can be present in its third dimension and in such case more empirical design guidelines should be provided to secure the use of visual variable on the three dimensional cartographic text and to provide a preferable and high efficiency design.

The dissertation focussed on two of the important users' characteristics, which are users' gender and expertise. Although they show only few significant differences in the between-group analysis, we think of between-group analysis as an essential analysis of map usability study. Another important users' characteristic is users' age, which could have a significant influence on both users' preference and efficiency of the cartographic text design.

Normally usability is measured by efficiency, effectiveness, and preference (ISO, 1994). The dissertation presented several usability studies that considered two aspects of usability, preference and efficiency. The effectiveness aspect was discussed only in Chapter 8. This issue left the door open for future research to study users' effectiveness regarding the application of size, shape, orientation and texture on the cartographic text.

It is crucial to learn how map users' search on the map. Such knowledge can enhance both map design and its usability. To reach this aim, eye tracking technique and software packages can be used to analyse users' search strategies. A good example of these software packages is the Visual Analytics Toolkit (Andrienko *et al.,* 2012),where eye movement data can be processed using methods like spatial generalization, adjustment of time reference, and spatio- temporal aggregation.

9.4 References

Andrienko, G., Andrienko, N., Burch, M., & Weiskopf, D. 2012. Visual Analytics Methodology for Eye Movement Studies. IEEE Transaction on Visualization and computer Graphics, 18(2), pp.2889-2898.

Andrienko, G., Andrienko, N., Rinzivillo, S., Nanni, M., Pedreschi, D., & Giannotti, F. 2009. Interactive Visual Clustering of Large Collections of Trajectories. IEEE VAST 2009,pp.3-10,, Atlanta City, New Jersey, USA.

Arntson, A. E. 2011. Graphic design basics. 5th edition, Thomson Wadsworth, Boston.

Bertin, J.1967. Semiology of Graphics: Diagram, Networks, Maps, Translated by William J. Berg, 2011, Esri press, California.

Bartz, B.S. 1970. An Analysis of the Typographic Legibility Literature, Assessment of Its Applicability to Cartography, The Cartographic Journal, 7(1), pp.6-17.

Brewer C.A. 1996. Guidelines for Selecting Colors for Diverging Schemes on Maps. The Cartographic Journal, 33(2), pp79-86.

Carter, R. C., & Silverstein, L. D. 2010. Size matters: Improved colour difference estimation for small visual targets. Journal of the Society for Information Display, 18(1), pp.17-28.

Carter, R., & Huertas, R. 2010. Ultra-Large Color Difference and Small Subtense. Color research and application, 35(1),pp.4-17.

Christ,R. E. 1975. Review and analysis of color coding research for visual displays, Human Factors, 17 (6), pp.542-570.

Dumas, J. S., & Redish, J.C. 1999. A practical guide to usability testing. Intellect Books, Exeter.
Fairbairn, D.J. 1993. On The Nature of Cartographic Text, The Cartographic Journal, 30(2), pp.104-111.
Gerber, R. V. 1981. Competence and performance in cartographic language. The Cartographic Journal, 18(8), pp.104-111.
Hambrick, D. Z., & Engle, R. W. 2002. Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. Cognitive psychology, 44(4), pp.339-387.
Hecht, S. 1928. The relation between visual acuity and illumination. The Journal of general physiology, 11(3), pp.255-281.
ISO. 1994. ISO9241-11: Ergonomic requirements for office work with visual display terminals (VDTs), Part11: Guidance on Usability.
Kaplan, S. 1987. Aesthetics, affect, and cognition environmental preference from an evolutionary perspective, Environment and behavior, 19(1), pp.3-32.
Hambrick, D. Z., & Engle, R. W. 2002. Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is power hypothesis. Cognitive Psychology, 44, 33-387.
Kraak, M.J & Ormeling, F. 2010. Cartography. Visualization of Spatial Data, 3rd edition, Person Prentic Hall, Harlow.
Ling, J. & van Schaik, P. 2002. The effect of text and background colour on visual search of Web pages. Displays, 23(5), pp.223-230.
Lliinsky, N. 2010. 'On Beauty'. In: Beautiful Visualization: Looking at Data Through the Eyes of Experts, ed. by Steele, J. & Lliinsky, N., Chapter 1, pp.1-13, O'Reilly, Canada.
MacEachren A.M. 1995. How Maps Work. Representation, Visualization, and Design, Guilford Press, NY.
MacEachren, A. M. 1982. The Role of Complexity and Sympolization Method in Thematic Map Effectiveness, Annals of the association of American Geographers, 72(4), pp.459-513.
memory capacity, and age on cognitive performance: An investigation of the
Murch, G. 1985. Colour graphics-blessing or ballyhoo?. Computer Graphics Forum. 4(2), pp.127-135.
Nielsen, J. 1993. Usability Engineering, Academia Press, San Francisco.
Ooms, K., De Maeyer, P., Fack, V., Van Assche, E., & Witlox, F. 2012. Interpreting maps through the eyes of expert and novice users. International Journal of Geographical Information Science, 26(10) ,pp.1773-1788.
Ooms, K., De Maeyer, P., & Fack, V. In press. Listen to the Map User: Cognition, Memory, and Expertise. The Cartographic Journal. DOI: http://dx.doi.org/10.1179/1743277413Y.0000000068.
Phillips, R. J., Noyes, E., & Audley, R. J. 1978. Searching for names on maps. The Cartographic Journal, 15(2), pp.72-77.
Phillips, R.J. 1981.Search for Target in a Random Arrangement of Names: an Eye Fixation Analysis, Canadian Journal of Psychology, 35(4), pp.330-346.
Popelka, S. & Brychtova, A. 2013. Eye-tracking study on different perception of 2D and 3D terrain visualisation. The Cartographic Journal, 50(3), pp.240-246.
Robinson, A. H. 1952. The Look of Maps: An Examination of Cartographic Design. University of Wisconsin Press.
Rubin, J. & Chisnell, D. 2008. Handbook of Usability Testing. How to Plan, Design and Conduct Effective Tests, 2nd edition, Wiley Publishing, Indianapolis.
Schenkman, B. N. & Jonsson, F. U. 2000. Aesthetics and preferences of web pages, Behaviour and Information Technology, 19(5), pp.367-377.

Shieh, K-K., Chen, M-T., & Chuang, J-H. 1997. Effect of Color Combination and Typography on Identification of Characters Briefly Presented on VDTs. International Journal of Human-Computer Interaction, 9(2), pp.169-181.

Shirreffs, W. S. 1992. Typography and the alphabet. The Cartographic Journal, 30(2), pp.97-103.

Shlaer, S. 1937. The relation between visual acuity and illumination. The Journal of general physiology, 21(2), pp.165-188.

Slocum.T.R, McMaster.R.B, Kessler.E.C, & Haward.H.H. 2005. Thematic Cartography and Geographic Visualization, 2nd edition, Person Prentic Hall, Upper Saddle River, NJ.

Tyner, J. A. 2010. Principles of Map Design, The Guilford press, NY.

Watzman, S. 2002. Visual design principles for usable interfaces. In The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications, pp.268-285.

Withycombe, J. G. 1929. Lettering on maps. The Geographical Journal, 73(5), pp.429-435.

Wolfe, J. M. 1994. Guided search 2.0: A revised model of visual search. Psychonomic Bulletin & Review, 1(2), pp.202-238.

10

Conclusions

This dissertation was based on the application of Bertin's visual variables on cartographic text design. To study how Bertin's variables influence the usability of cartographic text, cartographic text was considered as a symbol type. A series of usability studies was presented, using three different methodologies to evaluate digital maps and more specifically label design. First, usability of cartographic text design was investigated through a questionnaire that aimed to evaluate users' preference of different design. Second, we used reaction time measurements (RTM) to evaluate users' efficiency of certain designs, and third the eye tracking technique was used to evaluate both users' efficiency and effectiveness. Throughout the chapters, the 'power' of Bertins' variables, when used for cartographic text design, was described and the results of our experiments help setting a guideline for using size, shape, orientation, texture, and colour hue in cartographic text design.

During these studies, two types of data were collected. The first type of data is qualitative in nature, and was collected through questionnaires and eye tracking technique. The second type of data is more quantitative and was collected through RTM and eye tracking technique. In most cases, data was quantified and statistically analysed. However, the nature of the qualitative data was also used to describe user groups and other individuals' characteristics.

The focus of this dissertation is on how we can improve screen map usability by applying design variables on cartographic text. In correspondence with the first research objective, chapter 2 and chapter 4 explains how we can apply visual variables on label to improve both their preference and efficiency. Because text is always linked to language and more accurately to a lettering system, the dissertation considered cross-cultural comparisons of how we can use Bertin's variable to improve labelling. Therefore, four of the most used lettering systems (Latin, Cyrillic, Arabic, and Chinese) were implemented in corresponding experiments to study users' preference. Also, Latin, Arabic, and Chinese lettering systems were compared to study users' efficiency. The cross-cultural comparisons presented an overview of both users' preference and efficiency regarding the lettering systems and the main issues that can conflict both monolingual and multilingual mapping. Regarding the combination of different lettering system on maps and how they can be presented, multilingual mapping still require much consideration. The comparison demonstrated that each lettering system has an impact onto the typographic text design. The use of design parameters (based on Bertin's variables) cannot be applied similarly on different lettering systems.

Eye movement data presented a complete explanation of the effect of ground-figure design regarding labelling colours. Colours' foreground- background interaction was clarified by studying complementary colour influence on reading labels and comparing coloured labels with black labels. Reading coloured labels compared to reading black label was not significantly different apart from the case when colour rose is in the foreground and its spring green is in the background, in such case, using black label is more efficient than using the complementary colour. However, the research showed that more aspect of colour definition are equal in importance and shall be considered when studying colours such as colour difference and luminance contrast.

Although coloured maps are preferred over monochromic maps, colour design on maps is a controversial issue. Preference data, obtained in chapter 3, demonstrated the relationship between the cartographic text design and map background design. We found that the map background colour does not influence users' preference of cartographic text design, when labels are depicted in black. Additionally, when labels are depicted in colour and placed over a coloured background, users' preference of such design does not influence their efficiency. Therefore, to produce attractive and proffered maps, map designers should not adapt their textual design in regard to the maps' background colours.

To learn how users' search on maps, chapter 7 was devoted to meet this target. When considering the use of multi-packages software, eye tracking data can be well implemented and they can be analysed both quantitatively and qualitatively. The presented analysis of scanpaths, using OGAMA software, is influenced by the design and number of AOIs. The increase of AOIs numbers will refine the analysis and the design can modify all the corresponding analysis.

This presented work aims to contribute to the understanding of how map users respond to the functional designs of cartographic text on screen maps. This information is essential to create guidelines for the map designers regarding the map users' characteristics. The presented user studies show only few significant differences between users group, when measuring preference, efficiency and effectiveness. These measurements demonstrate that map designers should not adapt their typographic design in regard to their target audience. However, the results of our experiments trigger further research into the investigation of other characteristics of map users, such as age and cultural background, which may have a significant influence on the usability of cartographic text design.

11

Summaries

11.1 Summary

Cartographers employ their knowledge and experience to transfer the physical world into visual representation by using the cartographic language. The map user decodes the cartographic language according to his cartographic knowledge and recreates and interprets the depicted reality. Many influencing factors can guide this process such as knowledge, experience, psychological processes, abilities, aims, needs, interests, tasks, skills and other external conditions. Therefore, the created reality at the users' side will not match the reality created initially by the cartographer. This fact causes errors in the communication process (Kolácný, 1969).

It is likely that a 'bad' designed map leads to wrong communication at the users' side, and the 'well' designed map results in better communication with users. However, the attribute 'bad' and 'good' map design is not only determined by cartographic theoretician and designers but also by map users (Board and Taylor, 1977; Kraak and Ormeling, 2010).

Bertin's(1967) contribution to the graphic world is one of the earliest and the most important as it is considered the basic root for sign systems and design studies. Furthermore, it has been the main reference for cartographer since then. This research will focus on what Bertin presented to define and develop graphics. Additionally, it will consider his variables' applications on cartographic text and their influence on both users' preference and efficiency.

Bertin (1967) stated in the beginning of his book semiology of graphic: 'Graphic representation constitutes one of the basic sign-systems conceived by the human mind for the purpose of storing, understanding, and communicating essential information. As a "language" for the eye, graphics benefits from the ubiquitous properties of visual perception'. The statement focuses on two basic elements:

- how the sign system is organized, perceived, interpreted, and comprehended by the human mind. The information's level of organization within our mind is linked to data type and classification, which can be either quantitative or qualitative
- visual variables (at a fixed X,Y position, Bertin's variables are: size; shape, colour-hue-, value, orientation, and texture) and their properties' influence on visual perception: associativity, selectivity, order and quantity.

The evolution of the digital geospatial data handling has gained its momentum in the 1980s. Since then, more and more software packages were presented to cartography, for the purpose of both the production and the use of maps. Thus, two types of media are available to the users regarding the presentation of maps: paper and digital. Paper maps play (in the digital era) the traditional role of providing an overview of the geospatial relationships and pattern, meanwhile, the digital maps can also function as interface or index for extra information (Kraak, 2001). This dissertation will consider how Bertin's variable, which were proposed initially for semiotics of paper maps, can be applied to improve digital map design.

Fairbairn (1993) explored the functionality of the cartographic text and invited cartographers to consider the cartographic text as the fourth symbol type (in addition to point, line, and polygon). Treating map texts' as a symbol invites cartographer to consider using the visual variables to present different characteristics of the cartographic text.

The production of 'good' cartographic texts which provides high level of communication, their visualization and thus their design should consider perception and legibility, balance and hierarchy, discrimination, harmony, classification, transferability, and finally availability, cost, convenience, and suitability for reproduction (MacEachren, 1995; van den worm, 2001; Slocum *et al.,* 2005; Kraak and Ormeling, 2010; Tyner, 2010). The visualization of cartographic text should present the following requirements (Slocum *et al.,* 2005; Kraak and Ormeling, 2010; Tyner, 2010):

- they should present nominal differences between different data classes;
- they should be able to address hierarchy. Thus, cartographic text visualization is associated with differentiating between more and less important object or object categories);
- they should relate to all feature type (point, line and area).

The cartographic text is added to complete the visualization functions and to communicate map messages. Its function is always associated with the attributed data. Typographic variables like shape, colour hue, texture directionality, and orientation are effective to attribute qualitative data. While size, colour value, colour saturation, texture size, and texture density are effective to attribute quantitative data.

The fact that we cannot effectively read and interpret maps without their cartographic text (Fairbairn, 1993; Tyner, 2010), whose functions accomplish delivering map message, was the core motivation for this dissertation.

This dissertation comprises five research questions which are triggered by four research objectives:

- research objective A: improve the employment of visual variables in screen map design for better aesthetic and functional design;
- research objective B: investigate the influence of cartographic expertise and gender differences on map users' preferences and performance regarding cartographic text designs;
- research objective C: discuss the relationship between cartographic text colour and map background colour design; and
- research objective D: investigate the impact of lettering systems on cartographic text design.

These four main research objectives are decoded into five, more specific research questions:

- **RQ1**: How can the visual variables be applied to cartographic text, and more specifically map labels?
- **RQ2**: How do user characteristics influence the perception of label design?
- **RQ3**: How do lettering systems affect the usability of visual variables?
- **RQ4**: How do maps' background designs influence user preferences and performance together with the label design?
- **RQ5**: What type of search strategy do users follow when they search for labels over the map face and for certain map design conditions?

The description of the application of Bertin's visual variable onto cartographic text and the properties of these various typographic designs is presented in chapter 2. Additionally, the preference of two users' group (gender and expertise) towards the cartographic design was determined and the variations between the sub-groups (females versus males and experts versus novices) were analysed statistically, where few statistical differences between the subgroups preference wise were located.

Chapter 3 investigates the influence of maps' background colour onto users' preference of the cartographic text design. Four sets of experiments were designed simultaneously, each of which carries out the same typographic design, but with a basic change of map background colour design. Blank background map, grey scale map, warm colour map, and cold colour map were the four basic designs. The statistical analysis showed that no differences between users' preferences resulted from the change of map background colour design and thus no need to adapt the typographic design to the map colours.

Users' efficiency regarding cartographic text design is explained in detail in chapter 4. Two users group (gender and expertise) were involved in an online experiment which implemented a time measurement to evaluate users' efficiency. The efficiency of users' towards the cartographic text design was calculated when users locate target labels. Label designs have changed over the stimuli in correspondence to the application of visual variables onto the cartographic text.

In chapter 5 and 6 different lettering systems are implemented. Chapter 5 discuses users' preference regarding the application of Bertin's variables on four lettering system (Latin, Cyrillic, Arabic, and Chinese), and cultural comparison between the lettering systems was made. The comparisons between users' preference of Latin and Cyrillic designs showed identical users' preference of the different label designs, but the comparison of users' preference regarding Latin, Arabic, and Chinese showed significant differences of the different designs.

Chapter 6 discuses users' efficiency regarding the applications of visual variables on three lettering systems (Latin, Arabic, and Chinese). The chapter was designed for native users who were asked to locate a target label on a map. Users' efficiency towards label designs and different lettering systems was calculated and statistically compared. Users' efficiency deviated remarkably in correspondence with labels' design, which resulted from applying the visual variables on labels. Additionally, cultural comparison between the three lettering systems was made.

Chapter 7 focuses on the users' search strategies on maps (in the case of certain land use map). Therefore, an eye tracking experiment was conducted to demonstrate the users' search strategies and specially to look for strategic' clusters. This study engaged several open source software packages to analyse scan paths. Within this chapter, we continue to analyse the differences between users' groups as it distinguishes between the performance of the expertise and gender groups. When search tracks (scanpaths) were analysed and demonstrated, few search clusters appeared.

The visual variable colour was the focal point in chapter 8. Coloured labels in the foreground were studied in comparison to black labels. In addition to that, the interaction between colours in the foreground and their complementary in the background was analysed. Eye tracking data was used to interpret the users' efficiency and effectiveness. When data was collected, we distinguish between two users' groups' characteristics (expertise and gender)

In this dissertation, the application of visual variables was implemented to set the design of cartographic text. Although the sharp angles, firm edges, and complex curves of letters' shapes make them one of the most complicated visual elements, much of this complexity goes unnoticed because the shapes of letters are well known to the readers (Robinson, 1952). Cartographers acknowledged that the cartographic text shape is infinitive (Robinson, 1952; Bartz, 1970; MacEachren, 1995; Tyner, 2010). Yet, dividing the shapes of cartographic text into classes or categories, considering the fonts' essential characteristics, facilitates map depiction and uses the 'optimal' shapes for different functions. Different shapes of cartographic text have the characteristics that depict features which may either harmonize well or contrast well with the characteristic of feature shapes on a map.

Notwithstanding, in cartography the choice between upright (straight) and italic is usually based solely on the desire to differentiate between categories of data by change in type face. Yet, italic type is generally undesirable except for emphasis (Robinson 1952). This research is in coherence with Robinsons' description of italic cartographic text. Although he described it for paper mapping, the results agree that users are in favour of using straight cartographic text more than using italic text, even when using screen maps.

Our results showed positive responses towards the horizontal orientation of labels. This finding contributes to Robinsons' observations (1952) that when the letters are presented on a non-horizontal base they will attract attention by their inharmonious position and appearance. Although non-horizontal positioning of letters attracts attention, so many non-horizontal labels distract attention and this issue could have caused a delay of response time measurement in relation to horizontally placed labels. The study showed that users are in favour of horizontal orientation more than other sort of label placements (tilted orientation, i.e., along with the diagonal of the polygons' shape or mixed orientation, i.e., combines between horizontal and tilted).

The texture of cartographic text (labels) can be built by the combination of two or more visual variables. Therefore, when designed functionally, the cartographic texts help to correctly perceive the attributed features. The complexity of cartographic text design is most functional with texture design. The tested designs of labels' textures showed less preference and less efficiency with more complex labels' texture. Research on symbolization design complexity agrees with our finding considering memory of general pattern, and communicating general patterns (MacEachren, 1982).

Today, with the expanding use of screen maps and the considerably cheap printing costs, coloured maps have become ubiquitous. Colours allow great flexibility in map design as they facilitate the distinction between figure – ground elements (Tyner, 2010). Although the use of colour is considered desirable, we must be aware that its implications are challenging (e.g., later inhabitation explained by MacEachren, 1995). Colour can be used to enhance the effectiveness of graphical display. Additionally, colours are more preferable over black and white designs and while presenting beauty, colour is one of the strongest components in the preference space. Yet, the use of colour in visual display should not only be influenced by the subjective preference of the designer or users, but also by the cognitive perceptual constrains of the users. When colours are used properly, they can lead to faster search times (Christ, 1975; Murch, 1985; Kaplan, 1987; Brewer, 1996; Schenkman and Jonsson, 2000; Ling and van Schaik, 2002)

In contrary to Bartz (1970), who suggested using the typographic literature, which studied the typographic design on plain text, for cartographic text design to improve maps' typography, this research provided empirical confirmations that cartographic text has their own rules of design to provide efficient and preferred cartographic text design. Yet, some typographic design principles on plain text are still applicable on cartographic text such as the findings of Shieh *et al.* (1997). They demonstrated how type face significantly affected visual performance, as participants performed better with computer type than with the aesthetically more pleasing Kia type.

The individual's performance, in cartographic language can be affected by its own characteristics including intelligence, visual perception, map-reading ability, drawing ability, home environment, previous experiences with maps and attitudes toward maps (Gerber, 1981). Map users' characteristics influence the overall responses to special tasks (Nielsen, 1993; Dumas and Redish, 1999; Rubin and Chisnell, 2008, andothers).Yet, few significant differences were located in this research.

Although cartographers argued considering the cartographic text as a symbol (e.g., Robinson,1952 ; Fairbairn, 1993), we addressed the four distinguishable lettering systems as symbols and simultaneously applied Bertin's variables on these symbols. When the designs of Latin and Cyrillic were compared preference wise in chapter 5, no significant differences were located and this could have resulted essentially because both lettering systems encompass similar characteristics. Therefore, the research continued to compare preference wise between three different lettering systems, which have different characteristics (Latin, Arabic, and Chinese), where significant differences were located regarding cartographic text design. Additionally, the research in chapter 6 was able to locate significant differences regarding the efficiency of users, when these lettering systems were tested.

Addressing the 'optimal' cartographic text design is determined by their presentation conditions, the presentation functionality (preference versus efficiency), the type of labelled data (point, linear, and polygon), and finally, the target audience. However, when considering how visual variables are applied on maps' textual elements, regarding cartographic text size, shape, orientation, texture, and colour hue, this application influenced both users' preference and performance of different textual elements' designs. Additionally, the application of Bertin's variables on different lettering system influences differently both users' preference and performance. These findings form guidelines, which invite cartographic designer to adapt while they are presenting monolingual or multilingual static screen maps. These guidelines presented the most preferred size, shape, orientation, texture and colour. Simultaneously, the most efficient size, shape, orientation, texture and colour, regarding our experiment conditions were presented.

References

Bartz, B.S. 1970. An Analysis of the Typographic Legibility Literature, Assessment of Its Applicability to Cartography, The Cartographic Journal, 7(1), pp. 6-17.

Bertin, J.1967. Semiology of Graphics: Diagram, Networks, Maps, Translated by William J. Berg, 2011, Esri press, California.

Board, C., & Taylor, R. M. 1977. Perception and maps: Human factors in map design and interpretation. Transactions of the Institute of British Geographers, 2 (1), pp.19-36.

Brewer, C.A. 1996. Guidelines for Selecting Colors for Diverging Schemes on Maps. The Cartographic Journal, 33(2), pp. 79-86.

Christ, R. E. 1975. Review and analysis of color coding research for visual displays, Human Factors, 17 (6), pp.542-570.

Dumas, J. S., & Redish, J.C. 1999. A practical guide to usability testing. Intellect Books, USA.

Fairbairn, D.J. 1993. On The Nature of Cartographic Text, The Cartographic Journal, 30(2), pp.104-111

Gerber, R. V. 1981. Competence and performance in cartographic language. The Cartographic Journal, 18(8), pp.104-111.

Kaplan, S. 1987. Aesthetics, affect, and cognition environmental preference from an evolutionary perspective, Environment and behavior, 19(1), pp.3-32.

Kolácný, A. 1969. Cartographic information - A fundamental concept and term in modern cartography. The Cartographic Journal, 6(1), pp.47-49.
Kraak, M- J. 2001. 'Settings and Needs for Web Cartography', In: Web Cartography, ed. by Kraak, M.J. and Brown, A. Chapter 1, pp. 1-7, Taylor & Francis, Great BritainLondon.
Kraak, M.J & Ormeling, F. 2010. Cartography. Visualization of Spatial Data, 3rd edition, Person Prentic Hall, Harlow.
Ling, J. & van Schaik, P. 2002. The effect of text and background colour on visual search of Web pages. Displays, 23(5), pp.223-230.
MacEachren, A. M. 1982 .The Role of Complexity and Sympolization Method in Thematic Map Effectiveness., Annals of the association of American Geographers, 72(4), pp. 459-513.
MacEachren,. A.M. 1995. How Maps Work, Representation, Visualization and Design, the Guilford press, New York.
Murch, G. 1985. Colour graphics-blessing or ballyhoo?. Computer Graphics Forum. 4(2), pp.127-135.
Nielsen, J. 1993. Usability Engineering, Academia Press, San Francisco.
Robinson, A. H. 1952. The look of maps; An examination of cartographic design. Madison: University of Wisconsin Press.
Robinson, A. H., & Petchenik, B. B. 1975. The map as a communication system. The Cartographic Journal, 12(1), pp. 7-15.
Rubin, J. &Chisnell, D. 2008. Handbook of Usability Testing. How to Plan, Design and Conduct Effective Tests, 2nd edition, Wiley Publishing, Indianapolis.
Schenkman, B. N. & Jonsson, F. U. 2000. Aesthetics and preferences of web pages, Behaviour and Information Technology, 19(5), pp.367-377.
Shieh, K-K., Chen, M-T., & Chuang, J-H. 1997. Effect of Color Combination and Typography on Identification of Characters Briefly Presented on VDTs. International Journal of Human-Computer Interaction, 9(2), pp. 169-181.
Shirreffs, W. S. 1992. Typography and the alphabet. The Cartographic Journal, 30(2), pp. 97-103.
Slocum, T. R., McMaster, R. B., Kessler, E. C., & Haward, H. H. 2005. Thematic Cartography and Geographic Visualization, 2nd edition, Person Prentic Hall, Upper Saddle River, NJ.
Tyner, J. A. 2010. Principles of Map Design, The Guilford press, New York.
van den Worm, J. 2001. 'Web Map Design in Practice', In: Web Cartography, ed. by Kraak, M.J. and Brown, A. Chapter 7, pp. 87-107, Taylor & Francis, Great BritainLondon.

11.2 Samenvatting

Cartografen wenden hun kennis en ervaring aan om de fysieke wereld (visueel) af te beelden door gebruik te maken van de cartografische taal. De kaartgebruiker decodeert de cartografische taal in overeenstemming met zijn/haar cartografische kennis, herwerkt en interpreteert vervolgens de voorgestelde realiteit. Veel beïnvloedende factoren sturen dit proces, namelijk: kennis, ervaring, psychologische processen, vaardigheden, doelen, noden, interesses, taken, talenten en andere externe omstandigheden. Vandaar zal de gecreëerde realiteit (vanuit een gebruikersoogpunt) niet overeenstemmen met de initieel gevormde realiteit door de cartograaf. Dit feit veroorzaakt fouten in het communicatieproces (Kolácný, 1969).

Waarschijnlijk leidt een ‘slecht’ opgemaakte kaart tot slechte communicatie (langs gebruikerszijde) en de ‘goed’ opgemaakte kaart geeft betere resultaten in de communicatie met gebruikers. Desalniettemin, het label ‘slechte’ en ‘goede’ kaartopmaak wordt niet enkel bepaald door cartografische theoretici en ontwerpers maar ook door kaartgebruikers (Board and Taylor, 1977; Kraak & Ormeling, 2010).

Bertins (1967) bijdrage tot de grafische wereld is één van de vroegste en meest belangrijke aangezien het beschouwd wordt als de basis voor symboolsystemen en ontwerpstudies. Het is dan ook de belangrijkste referentie voor cartografen sindsdien. Dit onderzoek richt zich op wat Bertin vooropstelde qua definities en grafische ontwikkelingen. Bovendien neemt het de toepassing van zijn grafische variabelen op cartografische tekst en hun invloed ervan op zowel gebruikersvoorkeuren en efficiëntie in beschouwing.

Bertin stelde in het begin van zijn boek omtrent grafische semiologie: “grafische voorstellingen omvatten één van de basis tekensystemen bedacht door het menselijk brein met het doel essentiële informatie te bewaren, te begrijpen en te communiceren. Als een ‘taal’ voor het menselijk oog, genieten grafische data van de alomtegenwoordige eigenschappen van visuele perceptie”. De uiteenzetting richt zich op twee basiselementen:

- hoe het symboolsysteem is georganiseerd, ervaren, geïnterpreteerd en begrepen door de menselijke geest. Het organisatorisch informatieniveau in ons brein is gelinkt aan gegevenstypes en classificaties, wat kwantitatief of kwalitatief kan zijn;
- visuele variabelen (op een vaste X, Y positie) omvatten volgens Bertin: grootte, vorm, kleur, grijswaarde, oriëntatie en textuur. Daarnaast beschrijft hij de invloed van hun eigenschappen op visuele perceptie: associatie, selectiviteit, orde en hoeveelheid.

De evolutie van de digitale geospatiale gegevensbewerking bereikte zijn hoogtepunt in de jaren 1980. Sindsdien werden meer en meer softwarepakketten voorgesteld aan de cartografie met de intentie van zowel productie als kaartgebruik. Er zijn twee presentatiemedia beschikbaar voor het afbeelden van kaarten: digitale en papieren media. De papieren kaarten hebben (in het digitale tijdperk) een traditionele rol te vervullen: een overzicht bieden van geospatiale relaties en patronen. Ondertussen fungeren digitale versies eveneens als interface of als een index voor extra informatie (Kraak, 2001). Deze verhandeling zal bekijken op welke wijze Bertins variabelen, die initieel voorgesteld werden voor een semiologie van papieren kaarten, kan worden toegepast om het ontwerp van digitale kaarten te verbeteren.

Fairbairn (1993) verkende de functionaliteit van cartografische tekst(en) en zette de cartografen ertoe aan om de cartografische tekst als het vierde symbooltype te aanvaarden (als toevoeging aan punten, lijnen en polygonen). De behandeling van kaartteksten als een symbolen nodigt cartografen ertoe uit van gebruikt te maken van de visuele variabelen om verschillende karakteristieken van cartografische tekst voor te stellen.

"Goede" cartografische teksten kan men herkennen door een hoog communicatieniveau. Hun afbeelding en tevens hun ontwerp-, dienen rekening te houden met perceptie en leesbaarheid, balans en hiërarchie, waarneming, harmonie, classificatie, overdraagbaarheid en tenslotte beschikbaarheid, kost, belang en geschiktheid voor reproductie (MacEachren, 1995; van den worm, 2001; Slocum *et al.,* 2005; Kraak & Ormeling, 2010; Tyner, 2010).

De visualisatie van cartografische tekst zou moeten beantwoorden aan de volgende vereisten (Slocum *et al.,* 2005; Kraak & Ormeling, 2010; Tyner, 2010):

- ze zouden nominale verschillen tussen meerdere gegevensklassen moeten weergeven;
- ze zouden hiërarchische verschillen tussen meerdere gegevensklassen moeten kunnen weergeven. Cartografische tekstvisualisatie is aldus geassocieerd met het maken van een onderscheid tussen belangrijk en minder belangrijk(e) objecten of (object)categorieën;
- ze houden in verband moeten kunnen worden gebracht met alle objecttypes: punten, lijnen en vlakken.

De cartografische tekst is toegevoegd om de visualisatiefuncties te vervolledigen en om kaartboodschappen door te geven. De functie ervan houdt verband met de attribuutgegevens. Typografische variabelen zoals vorm, kleur, gerichtheid van textuur en oriëntatie zijn effectief om bij kwalitatieve data attribuutinformatie voor te stellen. Terwijl grootte, kleur-(waarde), kleurintensiteit, textuurgrootte en -dichtheid effectief zijn om bij kwantitatieve data attribuutinformatie voor te stellen.

Het feit dat we geen kaarten kunnen lezen en interpreteren zonder hun cartografische tekst (Fairbairn, 1993; Tyner, 2010), met al s functie het afleveren van de cartografische informatie, was de kernmotivatie voor deze verhandeling.

Dit proefschrift omvat 5 onderzoeksvragen die gegenereerd worden door vier onderzoeks-doelstellingen:

- onderzoeksdoelstelling A: het verbeteren van het gebruik van visuele variabelen bij het ontwerp van (scherm)kaart voor een beter esthetische en functionele design;
- onderzoeksdoelstelling B: het onderzoeken van de invloed van cartografische expertise- en genderverschillen op de voorkeuren van kaartgebruikers en hun prestaties m.b.t. het cartografisch tekstontwerp;
- onderzoeksdoelstelling C: het nagaan van de impact van het gebruik van een verschillend alfabet op cartografisch tekstontwerp;
- onderzoeksdoelstelling D: het behandelen van de relatie tussen de kleur van cartografische tekst en de achtergrondkleur van de kaart.

Deze 4 belangrijkste onderzoeksdoelstellingen worden bestudeerd aan de hand van 5 meer specifieke research vragen:

- RQ1: Hoe kunnen de visuele variabelen toegepast worden op de cartografische tekst en meer specifiek de kaartlabel?
- RQ2: Hoe beïnvloeden de karakteristieken van de kaartgebruikers de perceptie van het labelontwerp ?
- RQ3: Op welke manier heeft het alfabet een impact op de bruikbaarheid van visuele variabelen ?
- RQ4: Hoe beïnvloedt het achtergrondontwerp van kaarten de gebruikersvoorkeur en de prestaties aangaande het labelontwerp ?
- RQ5: Welke onderzoeksstrategie volgen gebruikers wanneer ze zoeken naar labels op het kaartbeeld en voor bepaalde kaartontwerpen?

De beschrijving van de toepassing van Bertin's visuele variabelen op cartografische tekst en de eigenschappen van deze verscheidene typografische ontwerpen wordt besproken in hoofdstuk 2. Aanvullend werden volgende gegevens statistisch geanalyseerd: de voorkeur van 2 gebruikersgroepen (gender en expertise) ten opzichte van het cartografisch ontwerp en de variatie tussen de subgroepen (vrouwen versus mannen en experten versus leken). De statistische analyse duidt aan dat weinig verschil is tussen de subgroepen qua voorkeur.

Hoofdstuk 3 onderzoekt grondig het effect van het achtergrondkleur op kaarten op de gebruikersvoorkeur voor een cartografisch tekstontwerp. Vier experimenten werden parallel ontworpen. Elk van hen voerde hetzelfde typografisch ontwerp uit maar met een verandering in het ontwerp van de kaartachtergrondkleur. De statistische analyse toonde geen verschil aan qua gebruikersvoorkeur met betrekking tot een gewijzigde achtergrondkleur en er is dus geen behoefte om het typografisch ontwerp aan de kaartkleuren aan te passen.

In hoofdstuk 4 wordt de gebruikersefficiëntie van cartografisch tekstontwerp in detail uitgelegd. Twee gebruikersgroepen (gender en expertise) waren betrokken in een online experiment dat gebruik maakte van tijdsmetingtechnieken om gebruikersefficiëntie te beoordelen. De efficiëntie van gebruikers tegenover cartografisch tekstontwerp werd berekend wanneer gebruikers specifieke labels moesten lokaliseren. Variaties in de visuele variabelen waren toegepast betreffende de cartografische tekst.

Hoofdstuk 5 en 6 focussen op een ander aspect van cartografische tekst, namelijk het gebruik van verschillende beletteringsystemen of alfabetten. Hoofdstuk 5 bestudeert de toepassing van Bertins variabelen op vier 'alfabetten' (het Latijnse, Cyrillische, Arabische en Chinese) en culturele verschillen tussen de systemen worden vergeleken. Na het onderzoeken van gebruikersvoorkeuren Latijnse en Cyrillische welke een gelijkaardige gebruikersvoorkeur aantoonden, toonden de studie van tekstlabels tussen Latijnse, Arabische en Chinese systemen duidelijk verschillen aan.

Hoofdstuk 6 bestudeert de gebruikersefficiëntie voor wat betreft de toepassing van visuele variabelen in drie beletteringsystemen (Latijnse, Arabische en Chinese). Deze studie werd ontworpen voor testen met gebruikers die het voorgestelde alfabet in hun moedertaal gebruiken. De gebruikersefficiëntie met betrekking tot labelontwerpen en verschillende alfabetsystemen werd berekend en statistisch vergeleken. De gebruikersefficiëntie week opvallend af overeenkomstig het labelontwerp, wat enerzijds bleek in de toepassing van visuele variabelen op labels. Anderzijds werd er een significant onderscheid opgemerkt tussen de gebruikersefficiëntie in verschillende alfabetsoorten.

Hoofdstuk 7 richt zich op de manier waarop gebruikers op kaarten zoeken (in het geval van een bepaalde kaart betreffende landgebruik). Hiertoe werd een oogbeweging (eye tracking) experiment opgezet om zoekstrategieën te demonstreren en vooral om uit te kijken naar strategische clusters. Verscheidene open source softwarepakketen werd gebruikt en de scanpaden werden onderzocht. In dit hoofdstuk wordt verder de verschillen tussen gebruikersgroepen ontleed aangezien het een onderscheid maakt tussen de expertise-uitslag en de gendergroepen. Zoekpatronen werden geanalyseerd en aangetoond, doch verschenen er weinig clusters.

De visuele variabele kleur domineerde hoofdstuk 8. Er werd een studie gemaakt van gekleurde labels op de voorgrond en dit in vergelijking met zwarte labels. Bovendien werd er een analyse gemaakt van de interactie tussen kleuren op de voorgrond en hun complement op de achtergrond. Eye tracking gegevens werden aangewend om gebruikersefficiëntie en -effectiviteit te doorgronden. Bij de dataverzameling werd steeds een onderscheid gemaakt tussen twee groepen met verschillende eigenschappen, met name op basis van expertise en geslacht.

In dit proefschrift werd de toepassing van visuele variabelen geïmplementeerd als hulp bij het ontwerp van cartografische tekst. Alhoewel de scherpe hoeken, scherpe randen en complexe krommingen van lettervormen deze tot één van de meest gecompliceerde visuele elementen maken, blijft toch veel van deze complexiteit onopgemerkt omdat de lettervorm goed gekend is bij de lezers (Robinson, 1952). Cartografen erkennen dat de cartografische tekstvorm onbepaald is (Robinson, 1952; Bartz, 1970; MacEachren, 1995; Tyner, 2010). En toch maakt het volgende het afbeelden van een kaart makkelijker: de indeling van cartografische tekstvormen in klassen of categorieën en het nadenken over de essentiële eigenschappen van de lettersoorten. De optimale vorm wordt gebruikt voor verschillende functies. Verschillende cartografische tekstvormen hebben als eigenschap dat voorgestelde kenmerken, die eerder goed harmoniëren of goed contrasteren met de karakteristieke vormen op een kaart.

Desondanks is in het vakgebied cartografie de keuze tussen recht en cursief gewoonlijk enkel gebaseerd op de nood tot differentiatie tussen gegevenscategorieën door wijziging in het uitzicht. Het cursieve lettertype is over het al gemeen toch ongewenst om iets te benadrukken (Robinson 1952). Het onderzoek houdt verband met Robinsons beschrijving van cursieve cartografische tekst. Alhoewel hij het beschreef voor papieren kaarten, bevestigden de resultaten dat gebruikers de voorkeur geven aan het gebruik van rechte cartografische tekst in de plaats van cursieve tekst, zelfs bij het gebruik van schermkaarten.

Onze resultaten demonstreerden positieve reacties betreffende de horizontale labeloriëntatie. Deze conclusie draagt bij tot Robinsons (1952)waarnemingen, nl. dat wanneer de letters voorgesteld worden op een niet-horizontale basis, ze meer aandacht trekken door hun onharmonische positie en verschijning. Alhoewel de niet-horizontale letterpositionering aandacht trekt, leiden vele niet-horizontale labels de aandacht af en dit zou een vertraging in reactietijd kunnen veroorzaken in vergelijking met de horizontaal geplaatste labels. De studie bewees dat gebruikers horizontale oriëntatie verkiezen boven een ander soort van labelplaatsing (schuine oriëntatie, dat wil zeggen langs de diagonaal van de polygoonvorm of gemengde oriëntatie, d.w. z. combinaties van horizontale en schuine stand).

De textuur van cartografische tekst (labels) wordt gevormd door de combinatie van twee of meer visuele variabelen. Daartoe, indien functioneel ontworpen, fungeren de cartografische teksten als correcte waarneming van de toegewezen eigenschappen.

Het ingewikkelde aspect van het cartografische tekstontwerp is meest functioneel bij textuurontwerp. De geteste ontwerpen van labeltextuur demonstreerden een mindere voorkeur en efficiëntie bij meer ingewikkelde labeltextuur. Onderzoek betreffende de complexiteit van symboolontwerp bevestigt onze bevinding aangaande het geheugen van een algemeen patroon en communicerende algemene patronen (MacEachren, 1982).

Vandaag, met de uitbreiding van het gebruik van kaarten op schermen en de behoorlijk goedkope printkosten, zijn gekleurde kaarten alomtegenwoordig. Kleuren laten een grotere flexibiliteit toe in kaartontwerp aangezien zij het onderscheid tussen figuren en grondelementen duidelijker maken (Tyner, 2010). Alhoewel kleurgebruik aanbevolen wordt moeten we ons ervan bewust zijn dat de implicaties ervan een uitdaging vormen (bijvoorbeeld latere bewoning? uitgelegd door MacEachren, 1995). Kleur wordt aangewend om de effectiviteit van grafische weergave te verhogen. Bovendien zijn kleuren te verkiezen boven zwart-wit ontwerpen en esthetisch gezien is kleur één van de belangrijkste componenten in de voorkeursbeleving. Nochtans zou het kleurgebruik in de visuele weergave niet enkel beïnvloed mogen worden door de subjectieve voorkeur van de ontwerper of de gebruikers, maar ook door de cognitieve perceptie van de gebruikers. Als kleuren op de juiste manier aangewend worden kan er sneller op kaart gezocht worden (Christ, 1975; Murch, 1985; Kaplan, 1987; Brewer, 1996; Schenkman and Jonsson, 2000; Ling and van Schaik, 2002).

In tegenstelling tot Bartz (1970) - die voorstelde om de typografische literatuur te gebruiken, waarin het typografisch ontwerp voor vlakke tekst bestudeerd wordt om het kaartontwerp te verbeteren - voorzag huidig onderzoek de empirische bevestiging dat cartografische tekst zijn eigen ontwerpregels bezit om te voorzien in een efficiënt en verkiesbaar cartografisch tekstdesign. Toch zijn sommige typografische ontwerpprincipes (op vlakke tekst) nog steeds toepasbaar op cartografische tekst zoals de bevindingen van Shieh *et al.* (1997). Zij toonden aan op welke manier het soort uitzicht de visuele performantie beduidend beïnvloedde, aangezien deelnemers beter presteerden met het computer type in de plaats van het esthetisch mooiere Kia type.

De prestaties van individuen in het gebruik van de cartografische taal kan bepaald worden door zijn/haar eigen karaktereigenschappen inclusief intelligentie, visuele perceptie, kaartleestalent, tekentalent, thuisomgeving, vroegere ervaring met kaarten en houding tegenover kaarten (Gerber, 1981). De eigenschappen van de kaartgebruikers hebben een effect op de algemene reacties bij speciale opgaven (Nielsen, 1993; Dumas and Redish, 1999; Rubin and Chisnell, 2008, and others). En toch werden er belangrijke verschillen in dit onderzoeksgebied vastgesteld .

Alhoewel cartografen onderling debatteerden betreffende de cartografische tekst als een symbool (e.g., Robinson,1952 ; Fairbairn, 1993), bespraken we de 4 onderscheiden lettersystemen als symbolen en pasten dan ook analogisch Bertins variabelen aangaande deze tekens toe. Als we voorkeuren in de Latijnse en Cyrillische ontwerpen vergeleken in hoofdstuk 5 werd er geen significant onderscheid opgemerkt en dit was voornamelijk het gevolg van het feit dat beide alfabetsoorten vergelijkbare karakteristieken vertoonden. Bijgevolg werd het onderzoek verdergezet om drie verschillende soorten alfabet te vergelijken die verschillende eigenschappen hebben (Latijn, Arabisch en Chinees), met grote verschillen qua cartografisch tekstdesign. Aanvullend was het onderzoek (in hoofdstuk 6) in staat om belangrijke verschillen te detecteren betreffende gebruikersefficiëntie, wanneer deze lettersystemen getest werden.

Aangaande het "optimale" cartografische tekstontwerp: dit wordt bepaald door de voorstellingscondities, de presentatiefunctionaliteit (voorkeur versus efficiëntie), het soort gelabelde gegeven (punt, lijn en polygoon) en de beoogde doelgroep.

Wanneer we echter bekijken hoe visuele variabelen toegepast worden op tekstuele kaartelementen aangaande cartografische tekstgrootte, vorm, oriëntatie, textuur en kleurschakering, bemerken we hoe deze toepassing een effect heeft op zowel gebruikersvoorkeur en -prestatie als op het ontwerp van verscheidene tekstuele elementen. Aanvullend bemerken we dat de toepassing van Bertins variabelen op verschillende lettersystemen de gebruikersvoorkeur en -prestatie op een andere wijze beïnvloedt. Deze bevindingen vormen richtlijnen die de cartografisch ontwerper uitnodigen zich aan te passen, waar nodig, bij het grafisch voorstellen van ééntalige of meertalige statische digitale kaarten. Deze richtlijnen stellen de meest verkozen grootte, vorm, richting, textuur en kleur voor; terzelfdertijd ook de meeste efficiënte grootte, vorm, richting, textuur en kleur, binnen onze experimentele omgeving

Referenties

Bartz, B.S. 1970. An Analysis of the Typographic Legibility Literature, Assessment of Its Applicability to Cartography, The Cartographic Journal, 7(1), pp. 6-17.

Bertin, J.1967. Semiology of Graphics: Diagram, Networks, Maps, Translated by William J. Berg, 2011, Esri press, California.

Board, C., & Taylor, R. M. 1977. Perception and maps: Human factors in map design and interpretation. Transactions of the Institute of British Geographers, 2 (1), pp.19-36.

Brewer, C.A. 1996. Guidelines for Selecting Colors for Diverging Schemes on Maps. The Cartographic Journal, 33(2), pp. 79-86.

Christ, R. E. 1975. Review and analysis of color coding research for visual displays, Human Factors, 17 (6), pp.542-570.

Dumas, J. S., & Redish, J.C. 1999. A practical guide to usability testing. Intellect Books, USA.

Fairbairn, D.J. 1993. On The Nature of Cartographic Text, The Cartographic Journal, 30(2), pp.104-111

Gerber, R. V. 1981. Competence and performance in cartographic language. The Cartographic Journal, 18(8), pp.104-111.

Kaplan, S. 1987. Aesthetics, affect, and cognition environmental preference from an evolutionary perspective, Environment and behavior, 19(1), pp.3-32.

Kolácný, A. 1969. Cartographic information - A fundamental concept and term in modern cartography. The Cartographic Journal, 6(1), pp.47-49.

Kraak, M- J. 2001. 'Settings and Needs for Web Cartography', In: Web Cartography, ed. by Kraak, M.J. and Brown, A. Chapter 1, pp. 1-7, Taylor & Francis, Great BritainLondon.

Kraak, M.J & Ormeling, F. 2010. Cartography. Visualization of Spatial Data, 3rd edition, Person Prentic Hall, Harlow.

Ling, J. & van Schaik, P. 2002. The effect of text and background colour on visual search of Web pages. Displays, 23(5), pp.223-230.

MacEachren, A. M. 1982 .The Role of Complexity and Sympolization Method in Thematic Map Effectiveness., Annals of the association of American Geographers, 72(4), pp. 459-513.

MacEachren,. A.M. 1995. How Maps Work, Representation, Visualization and Design, the Guilford press, New York.

Murch, G. 1985. Colour graphics-blessing or ballyhoo?. Computer Graphics Forum. 4(2), pp.127-135.

Nielsen, J. 1993. Usability Engineering, Academia Press, San Francisco.

Robinson, A. H. 1952. The look of maps; An examination of cartographic design. Madison: University of Wisconsin Press.
Robinson, A. H., & Petchenik, B. B. 1975. The map as a communication system. The Cartographic Journal, 12(1), pp. 7-15.
Rubin, J. &Chisnell, D. 2008. Handbook of Usability Testing. How to Plan, Design and Conduct Effective Tests, 2nd edition, Wiley Publishing, Indianapolis.
Schenkman, B. N. & Jonsson, F. U. 2000. Aesthetics and preferences of web pages, Behaviour and Information Technology, 19(5), pp.367-377.
Shieh, K-K., Chen, M-T., & Chuang, J-H. 1997. Effect of Color Combination and Typography on Identification of Characters Briefly Presented on VDTs. International Journal of Human-Computer Interaction, 9(2), pp. 169-181.
Shirreffs, W. S. 1992. Typography and the alphabet. The Cartographic Journal, 30(2), pp. 97-103.
Slocum, T. R., McMaster, R. B., Kessler, E. C., & Haward, H. H. 2005. Thematic Cartography and Geographic Visualization, 2nd edition, Person Prentic Hall, Upper Saddle River, NJ.
Tyner, J. A. 2010. Principles of Map Design, The Guilford press, New York.
van den Worm, J. 2001. 'Web Map Design in Practice', In: Web Cartography, ed. by Kraak, M.J. and Brown, A. Chapter 7, pp. 87-107, Taylor & Francis, Great BritainLondon.

Questionnaires

Pre- test Questionnaire
Used in Chapter 2, chapter 3, chapter 4, chapter 5 and chapter 6.

1. What is your date of birth?

Day	Month	Year

2. Please do specify your gender.

Female	Male

3. Are you colour blinded?

Yes	No

4. What is your mother tongue?
Choose one of the following answers

Afrikaans	Georgian	Persian
Albanian	German	Polish
Arabic	Greek	Portuguese
Armenian	Haitian Creole	Romanian
Azerbaijani	Hebrew	Russian
Basque	Hindi	Serbian
Belarusian	Hungarian	Slovak
Bulgarian	Icelandic	Slovenian
Catalan	Indonesian	Spanish
Chinese	Irish	Swahili
Croatian	Italian	Swedish
Czech	Japanese	Thai
Danish	Korean	Turkish
Dutch	Latin	Ukrainian
English	Latvian	Urdu
Estonian	Lithuanian	Vietnamese
Filipino	Macedonian	Welsh
Finnish	Malay	Yiddish
French	Maltese	Other
Galician	Norwegian	Please specify

5. Could you please specify the languages you speak?
Check any that apply

Afrikaans	Georgian	Persian
Albanian	German	Polish
Arabic	Greek	Portuguese
Armenian	Haitian Creole	Romanian
Azerbaijani	Hebrew	Russian
Basque	Hindi	Serbian
Belarusian	Hungarian	Slovak
Bulgarian	Icelandic	Slovenian
Catalan	Indonesian	Spanish
Chinese	Irish	Swahili

Croatian	Italian	Swedish
Czech	Japanese	Thai
Danish	Korean	Turkish
Dutch	Latin	Ukrainian
English	Latvian	Urdu
Estonian	Lithuanian	Vietnamese
Filipino	Macedonian	Welsh
Finnish	Malay	Yiddish
French	Maltese	Other Please specify
Galician	Norwegian	

6. What is your current study?
(For example: First year Bachelor Geography)

7. What did you study last year?
(For example: First year Bachelor Geography)

8. What is your highest diploma so far?
Choose one of the following answers.

Secondary school	Bachelor	Master	PhD

9. How often do you use maps?
Choose one of the following answers

Daily	A few times a week	A few times a month	A few times a year

10. What type of maps do you usually use?
Choose one of the following answers

Internet maps
GPS maps
Atlases
Paper maps
Others

Post-test Questionnaire used in Chapter 8

1. What is your date of birth?

Day	Month	Year

2. Please do specify your gender.

Female	Male

3. Are you colour blinded?

Yes	No

4. What is your highest diploma so far?
Choose one of the following answers.

Secondary school	Bachelor	Master	PhD

5. How often do you use maps?
Choose one of the following answers

Daily	A few times a week	A few times a month	A few times a year

6. What type of maps do you usually use?
Choose one of the following answers

Internet maps
GPS maps
Atlases
Paper maps
Others

7. Which maps do you prefer more?

Maps with black labels	Maps with colour labels

8. Please specify the reasons for your preference.

Curriculium Vitae

Rasha deeb was born in Damascus (Syria) on the 28th of August in 1984. In 2002 she started her academic education at the Department of Geography (Damascus university). She obtained her Master's degree (*magna cum laude*) in Cartography, geographic information systems, and remote sensing application.

she worked as academic assistance in 2008-2009. As an assistance she was involved in several courses in the field of her major; Cartography, GIS, Computer Cartography, Map Design, Remote sensing, Image processing, historical cartography.

Rasha contributed to several international conferences, where her presentations were evaluated very positively by the scientific cartographic community. She is the author of a number of papers that are published in leading international journals in the field of cartography and GIS.

This dissertation presents a series of usability studies, which examines the usability of the application of visual variables on cartographic text. Labels' size, shape, orientation, texture, and colour were tested. The study also examines different lettering systems and their impact on cartographic text design. The obtained users' preference, time measurement, questionnaires and eye tracking data were analysed both qualitatively and quantitatively. Insights are acquired to improve the quality of map through Improving cartographic text design.