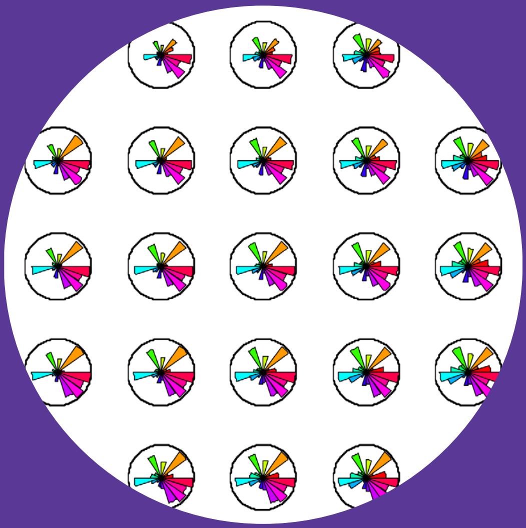
Department of Media Technology

Human abilities to perceive, understand, and manage multi-dimensional information with visualizations

Mikko Berg



Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 30/2012

Human abilities to perceive, understand, and manage multi-dimensional information with visualizations

Mikko Berg

Doctoral dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the School of Science for public examination and debate in Auditorium TUAS TU1 at the Aalto University (Espoo, Finland) on the 27th of April 2012 at 12.00.

Aalto University
School of Science
Department of Media Technology

Supervisor
Prof. Tapio Takala

Instructor
Ilpo Kojo

Preliminary examiners
Prof. Göte Nyman, University of Helsinki, Finland
Harri Siirtola, University of Tampere, Finland

Opponent
Prof. Erica de Vries, Université Pierre-Mendès-France, France

Aalto University publication series
DOCTORAL DISSERTATIONS 30/2012

ISBN 978-952-60-4548-1 (printed)
ISBN 978-952-60-4549-8 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Unigrafia Oy
Helsinki 2012

Finland

The dissertation can be read at http://lib.tkk.fi/Diss/

441 697
Printed matter

A! Aalto University

Abstract

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Author
Mikko Berg

Name of the doctoral dissertation
Human abilities to perceive, understand, and manage multi-dimensional information with visualizations

Publisher School of Science

Unit Department of Media Technology

Series Aalto University publication series DOCTORAL DISSERTATIONS 30/2012

Field of research Media Technology

Manuscript submitted 20 June 2011 **Manuscript revised** 12 October 2011

Date of the defence 27 April 2012 **Language** English

☐ **Monograph** ☒ **Article dissertation (summary + original articles)**

Abstract

Visualizations facilitate presenting information in a form adapted to human visual system. The thesis defines this adapting more precisely with a framework of human-visualization interaction. Human capacity limits are used to identify critical subtasks in interaction that specify whether the data quantities and complexity of relations are understandable. In reality, use context and personal characteristics also impact understanding. The framework identifies commonalities of human processing in context of large quantities of multi-dimensional data. The novel extension is to identify multi-dimensionality from tasks of ordinary people, such as shopping or voting. According to the hypothesis ordinary people can also understand more complex information with the help of visualized dimension-reduction algorithms (e.g., MDS or SOM). Such context and haste give prominence to perceptual and cognitive processes, the critical subtasks, and the related visual parameters. Unfortunately, the interactive effects of the related parameters on performance cannot be predicted with real applications, and thus applied studies are needed.

The proposed multi-disciplinary framework is based on reviewing empirical findings about humans. First, cognitive science about human concept formation is used to provide indications about how to pre-processes data into more readily understandable form. Second, vision research, experimental psychology, and neuroscience are used to explain the interaction, when information is received through vision and the input is influenced through gaze shifts and manual operations. Third, findings are reviewed about conditions, in which processing is likely to be externalized to visualization or tool as opposed to being performed by human mind. The framework is then applied for psychophysically controlled reaction time and eye movement experiments based on selective review of methods and critical visual parameters in research of visual searching. The evaluated tasks, searching and integrating, were chosen based on both behavioural studies with visualizations and implications of the framework. The ordinary people interacting with visualizations were evaluated indirectly from the popularity, feedback, and observations during two Finnish elections.

The framework led to three suggested principles. First, externalizing memory tasks and information processing to visualizations reduces mental load. Second, relations of dimensions can be uncovered by perceiving consequences of own manual actions. Third, peripheral vision is suitable for representing coarse outlines, for instance depicting dimension reduction. Facilitating rapid gaze shifts is important for all of these principles.

Keywords visualization, multi-dimensional, vision, perception, cognition, eye movements

ISBN (printed) 978-952-60-4548-1 **ISBN (pdf)** 978-952-60-4549-8

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Espoo **Location of printing** Helsinki **Year** 2012

Pages 175 **The dissertation can be read at** http://lib.tkk.fi/Diss/

Aalto-yliopisto

Tiivistelmä

Aalto-yliopisto, PL 11000, 00076 Aalto www.aalto.fi

Tekijä
Mikko Berg

Väitöskirjan nimi
Ihmisen kyky havaita, ymmärtää ja hallita moniulotteista informaatiota visualisointien avulla

Julkaisija Perustieteiden korkeakoulu

Yksikkö Mediatekniikan laitos

Sarja Aalto University publication series DOCTORAL DISSERTATIONS 30/2012

Tutkimusala Mediatekniikka

Käsikirjoituksen pvm 20.06.2011 **Korjatun käsikirjoituksen pvm** 12.10.2011

Väitöspäivä 27.04.2012 **Kieli** Englanti

☐ **Monografia** ☒ **Yhdistelmäväitöskirja (yhteenveto-osa + erillisartikkelit)**

Tiivistelmä

Visualisoinnit mahdollistavat informaation esittämisen näköjärjestelmää laajasti hyödyntävässä muodossa kuvina. Väitöskirjassa määritellään näköjärjestelmän hyödyntämistä aiempaa täsmällisemmin ihmisen ja visualisoinnin vuorovaikutuksen viitekehyksen avulla. Ihmisen suorituskyvyn rajoitteita tarkastellaan kriittisissä osatehtävissä aineiston kasvaessa ja sen sisäisten yhteyksien monimutkaistuessa. Todellisuudessa ymmärryskykyyn vaikuttavat myös aiheyhteys ja ihmisten erityispiirteet. Viitekehys kuvaa ihmisille yhteisiä tiedonkäsittelyn piirteitä suurien ja moniulotteisten aineistojen yhteydessä. Erityisesti tarkastellaan tavallisten ihmisten kohtaamaa moniulotteisuutta esimerkiksi ostotilanteissa tai äänestettäessä. Esitetyn hypoteesin mukaan myos he voivat hyötyä ulottuvuuksien redusoinnista (esim. MDS tai SOM). Nämä olosuhteet ja kiire korostavat havaitsemisen ja kognition kannalta kriittisten osatehtävien ja visuaalisten ominaisuuksien merkitystä. Näiden visuaalisten ominaisuuksien yhteisvaikutuksia suoriutumiseen ei valitettavasti voida kuitenkaan ennustaa tarkasti todellisissa sovelluksissa, ja siksi tarvitaankin soveltavaa tutkimusta.

Esitetty viitekehys perustuu poikkitieteellisen ihmistutkimuksen katsauksiin. Ensiksi, käsitteenmuodostukseen syventyvää kognitiotiedettä sovelletaan aineiston esikäsittelemiseksi ymmärrettävämpään muotoon. Toiseksi, viitekehykseen liittyy näkötutkimusta, kokeellista psykologiaa ja aivotutkimusta , kun informaatiota vastaanotetaan näönvaraisesti ja tähän syötteeseen vaikutetaan silmiä liikuttelemalla ja käsin toimimalla. Kolmanneksi, katsauksessa arvioidaan olosuhteita, joissa toimintoja ulkoistetaan visualisointiin tai työkaluun mielen sisäisten toimintojen sijaan. Lopuksi, viitekehystä sovelletaan reaktioaikoja ja silmanliikkeitä mittaavissa psykofysiikan kokeissa, joiden menetelmät perustuvat valikoituun katsaukseen visuaalisesta hausta ja sen kriittisista visuaalisista parametreista. Arvioitavat tehtävät, etsiminen ja yhdistely, valittiin visualisointien käytön tutkimusten ja esitetyn viitekehyksen perusteella. Tavallisten ihmisten visualisointien käyttöä arvioitiin epäsuorasti sovelluksen suosion, annetun palautteen ja havaintojen perusteella kaksissa suomalaisissa vaaleissa.

Viitekehyksestä johdettiin kolme periaatetta. (i) Muistin ja tiedonkäsittelyn ulkoistaminen visualisointeihin vähentää kognitiivista kuormitusta. (ii) Oman toiminnan seurauksien tarkkailu paljastaa ulottuvuuksien yhteyksiä. (iii) Ääreisnäkö soveltuu aineiston karkeille yleishahmoille, esimerkiksi esittämään ulottuvuuksien reduktiota. Nopean katseen kohdistamisen tukeminen on tärkeää kaikille näille periaatteille.

Avainsanat visualisointi, moniulotteisuus, näkö, havaitseminen, kognitio, silmänliikkeet

ISBN (painettu) 978-952-60-4548-1 **ISBN (pdf)** 978-952-60-4549-8

ISSN-L 1799-4934 **ISSN (painettu)** 1799-4934 **ISSN (pdf)** 1799-4942

Julkaisupaikka Espoo **Painopaikka** Helsinki **Vuosi** 2012

Sivumäärä 175 **Luettavissa verkossa osoitteessa** http://lib.tkk.fi/Diss/

Publications

The thesis consists of a summary (for which contents are outlined above) and the following publications:

[I] Berg, M., Kaipainen, M., & Kojo, I. (2004). "Enhancing usability of the similarity map for more accessible politics". Proceedings in 8th ERCIM workshop - User interfaces for all, Palais Eschenbach, Vienna, Austria, 39-46.

[II] Berg, M., Marttila, T., Kaipainen, M., & Kojo, I. (2006). "Exploring political agendas with advanced visualization and interface tools". e-Service Journal, 4(2), 47-63.

[III] Berg, M., Schleimer, J.-H., Särelä, J., & Honkela, T. (2005). "Category learning by formation of regions in conceptual spaces". In L. Magnani & R. Dossena (Eds.), Computation, Philosophy, and Cognition. (pp. 381-396). London, UK: College Publications

[IV] Berg, M. (2008). "Katseenliikkeet ja harha terävästi nähdystä näkökentästä". Psykologia, 3, 193-204. [Translated in Appendix A]

[V] Berg, M. & Kojo, I. (2012).“Integrating complex information with object displays: Psychophysical evaluation of outlines”. Behaviour & Information Technology, 31(2), 155-169.

[VI] Berg, M., Kojo, I., & Laarni, J. (2010). “Object displays for identifying multidimensional outliers from crowded visual periphery”. Journal of Visual Communication and Image Representation, 21, 880-888.

Publications and author's contribution

Included Publications fall into three general categories. First describes an application (1 & II), second present the underlying theories (III & IV), and third present empirical evaluations (V & VI). All Publications were based on similar findings and assumptions discussed in this summary. However, their specific objectives, formats and forums are different. The Publications I, V, and VI (also some in Publ. II) report new empirical information about the perception and use of visualization. Other Publications are more theoretical or discuss varying experimental findings reported elsewhere.

Publication I - Berg, M., Kaipainen, M., & Kojo, I. (2004). "Enhancing usability of the similarity map for more accessible politics". Proceedings in 8th ERCIM workshop - User interfaces for all, Palais Eschenbach, Vienna, Austria. (pp. 39-46).

The iterative design was described to improve the usability of SOM-based (Self-Organizing Maps, Sect. 2.3) similarity map prototypes. In a *similarity map*, more similar political candidates are placed closer together than dissimilar ones. The author designed and implemented several prototype interfaces, which were evaluated using interviews. The author was responsible for designing and conducting the interviews. At the time professor Mauri Kaipainen was supervising the author's master's thesis and provided background for the very early development of the very first functionalities of the application. The author provided the background for the usability literature.

Publication II - Berg, M., Marttila, T., Kaipainen, M., & Kojo, I. (2006). "Exploring political agendas with advanced visualization and interface tools". e-Service Journal, 4(2), 47-63.

The published service used in the nation-wide Finnish EU elections and communal elections in 2004 was described. In addition to the similarity map, the service included sector diagram representations, which were not evaluated in prototype development. Instead, the use of sector diagrams was later evaluated in psychophysical experiments in Publication V and VI. Publication II focuses on navigation, user observations, and user feedback. The author was responsible of conducting the prototype development to specify the visual layout and included functionalities of the application. Topias Marttila converted the Java versions of the interface to Flash platform before Alma media published the service. The idea to include also sector diagrams came from Ilpo Kojo. The author was responsible for discussing the usability of the application and implications of added exploration. The earlier election engines had only provided a list of best-matching candidates, but not explorative functions.

Publication III - Berg, M., Schleimer, J.-H., Särelä, J., & Honkela, T. (2005). "Category learning by formation of regions in conceptual spaces". In L. Magnani & R. Dossena (Eds.), Computation, Philosophy, and Cognition. (pp. 381-396). London, UK: College Publications

Gärdenfors' (2000) Conceptual Spaces theory was further developed by being more explicit on category learning process. The learned categories guide the understanding and subsequent decision-making when exploring visualized multi-dimensional data. The author was responsible for the theoretical discussion (on Conceptual Spaces, dynamical systems theory, and prototype theory) in the article. The theoretical coherence in the article was obtained through long discourse, where the author contributed most for the conceptual spaces theory and categorization literature. Jan-

Henrik Schleimer and Jaakko Särelä contributed for the Artificial Neural Network literature. Jan-Henrik Schleimer implemented the clustering algorithms and plotted the figures.

Publication V - Berg, M. & Kojo, I. (2012).“Integrating complex information with object displays: Psychophysical evaluation of outlines”. Behaviour & Information Technology, 31(2), 155-169

The perception of individual graphs was studied. The earlier psychophysical experiments with graphs are reviewed and two new experiments are reported. The author designed and conducted the experiments. The author also conducted the calculations of the necessary measures (similarity distances, staircase-algorithms etc.) and the statistical testing. As an experienced vision researcher, Ilpo Kojo, gave many advises to improve the visual and ergonomic conditions of the experiments. The author provided the review of psychophysical experiments with visualizations and other related literature.

Publication VI - Berg, M., Kojo, I., & Laarni, J. (2010). “Object displays for identifying multidimensional outliers from crowded visual periphery”. Journal of Visual Communication and Image Representation, 21, 880-888.

The earlier research related to eye movement and peripheral vision were summarized and four new visual search experiments were reported measuring reaction times and accuracy using either keyboard or eye movements. The objective was to determine whether the graphs are evaluated one at a time or all at once. The idea for the experiments 1 and 2 came from Jari Laarni, and the idea for the experiments 3 and 4 came from Ilpo Kojo. The author was responsible for reviewing the reference literature and analyzing the new results in terms of earlier ones. For instance, the author calculated eccentricities and distances to evaluate peripheral crowding as suggested in referenced literature.

Abbreviations

2-AFC	Two Alternative Forced Choice
CS	Conceptual Space(s)
DR	Dimension Reduction
fMRI	functional Magnetic Resonance Imaging
GUI	Graphical User Interface
HCI	Human-Computer Interaction
HUT	Helsinki University of Technology
JND	Just-Noticeable Difference
M-D	Multi-Dimensional
MDS	Multi-Dimensional Scaling
PCA	Principal Component Analysis
RT	Reaction Time
SOM	Self-Organizing Map
UI	User Interface
VCR	VideoCassetteRecord

Contents

1 INTRODUCTION

1.1 Background

Visualizations have already been used for few centuries to improve understanding of complex information. During the past few decades, computers have changed the visualization more than any other factor. The use of computers has increased the amount of represented data, its dimensionality, and the potential ways of user interaction. That way the new technology has also brought new challenges for the users of visualizations. Due to the new technology, the important relations in the represented data have become more complex. This complexity causes difficulties for all humans due to the limited capacities of visual-cognitive processes. These challenges are so important for users' understanding that the challenges can compromise all the benefits of using advanced computer-based visualizations. This is especially true for the largest potential user-group of people that are neither scientist nor experts with the given data. This thesis discusses those aspects of interactive visualization that are critical for understanding the visualized information. Special attention is paid to the initial phase, when the user explores complex data. First, the critical subtasks for this task are identified. The subtasks are then converted (operationalized) into criteria to compare and evaluate different graphs by experiments of cognitive psychology. The most relevant psychological experiments utilize quantitative methods of psychophysics. The suggested psychophysical criteria differ from those developed for static graphs in 1980s.

The thesis presents evidence for sequentiality or dynamics in perception of multi-dimensional visualizations. The presence of this sequentiality undermines the importance of the traditional psychophysical evaluation criteria. The sequentiality stems from the inability to receive all the represented data at once, when the quantities are numerous and the dimensionality is high. The importance of the sequentiality has increased with modern computer visualization as the size of easily available information has increased. Nevertheless, some data-intensive visualizations have existed already in 19th century (Tufte 1983), and all visualizations of the present day are certainly not data-intensive. However, many human decisions could be improved if more data could be taken into account. The suggested sequential framework to deliver information to the users is presented in Figure 1 and the included component parts are explained in the following. After defining the components, users' perceptual and cognitive challenges are outlined and related to the sequentiality in the framework. This constitutes the most important new contribution of this thesis to the framework.

Parts of the interactive framework

The first part consists of the given multi-dimensional data ['MD DATA' in Fig. 1]. This thesis will focus on the challenge of comparing similarities and finding best-matches from sets of discrete data entities with very many attributes, corresponding to standard numeric tables. The numeric table consist of discrete values, for which rows mark the data entities and columns mark the different variables that define the similarity between the entities. These variables are considered as dimensions when they are combined in abstract mathematical data space. Such was the case, when non-experienced users compared political opinions from hundreds of candidates and searched for the best alternative just before the communal elections of 2004 in Helsinki (Publ. II). The use of the visualization involved is analyzed in this thesis. The general reasons for creating such visualizations have been the challenges related to (i) comparing large number of items and (ii) utilizing large number of features (dimensions) in these comparisons (Berg & Kojo, 2011). The first factor in the election data resulted from the large number of candidates (hundreds) that needed to be compared. The second factor resulted from the different political issues that each candidate was commenting.

The second part in the typical visualization scheme is preprocessing of data before the visualization. Dimension reduction algorithms ['DR' in Fig. 1] are potentially a very useful preprocessing method to handle some of the challenges related to multi-dimensionality. Scientists and statisticians have used such algorithms as Principal Component Analysis (PCA, Pearson 1901), MultiDimensional Scaling (MDS, Shepard 1962), Self-Organizing feature Maps (SOM, Kohonen 1982; 2002) for quite some time to deal with large quantities and complexity of data. The comparison of different DR algorithms is outside the scope of the thesis, but the aim is to make their end-results more understandable to the common public. This is needed because general public is also shown to benefit from their use. DR algorithms can create summaries and simplifications and empower people to utilize more of the complex information available for their decisions.

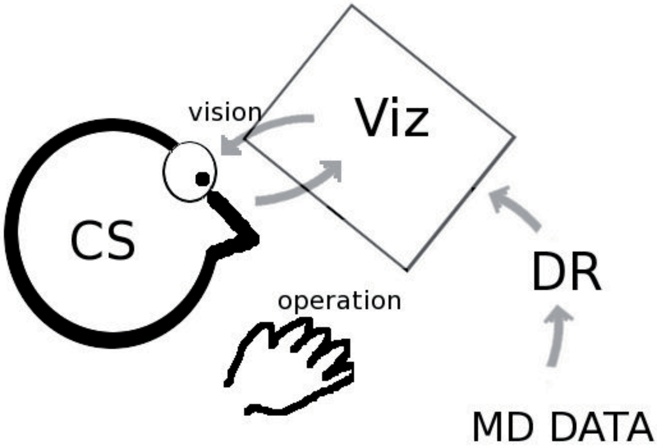


Fig. 1. Schematic illustration of studied interaction. The conceptualization process of multi-dimensional visualization includes the following parts: [MD DATA] represented multi-dimensional data, [DR] algorithmic dimension-reduction, [Viz] visualization, and [CS] human concept formation in Conceptual Spaces. The interaction between user and visualization includes: [vision] human vision that not only receives the input, but also influences the input with gaze shifts, and [operation] manual operations that influence the way visualizations are presented.

The problem is that the visualizations of dimension-reduction algorithms might not be readily accessible to general public, such as voters. Inexperienced users might not be able to evaluate the necessary questions of validity; the appropriateness of the algorithm for the given data and the possibility of losing important data in the simplification process. In the context of Internet-use, users should also be convinced that the information processing is informative, consistent, and relevant for their purposes. The offered solution is to convert validity assessments into visual assessments of similarity or consistency. The idea is that user could detect questionable dimension reduction from the discrepancies between visualizations of dimension-reduced data and raw data [‘MD DATA’ in Fig. 1]. Users can make better-informed decision if they can trust on visualizations of dimension-reduced data [‘Viz’], instead of being limited to their cognitive processing capacities [‘CS’].

Perceptual and cognitive challenges

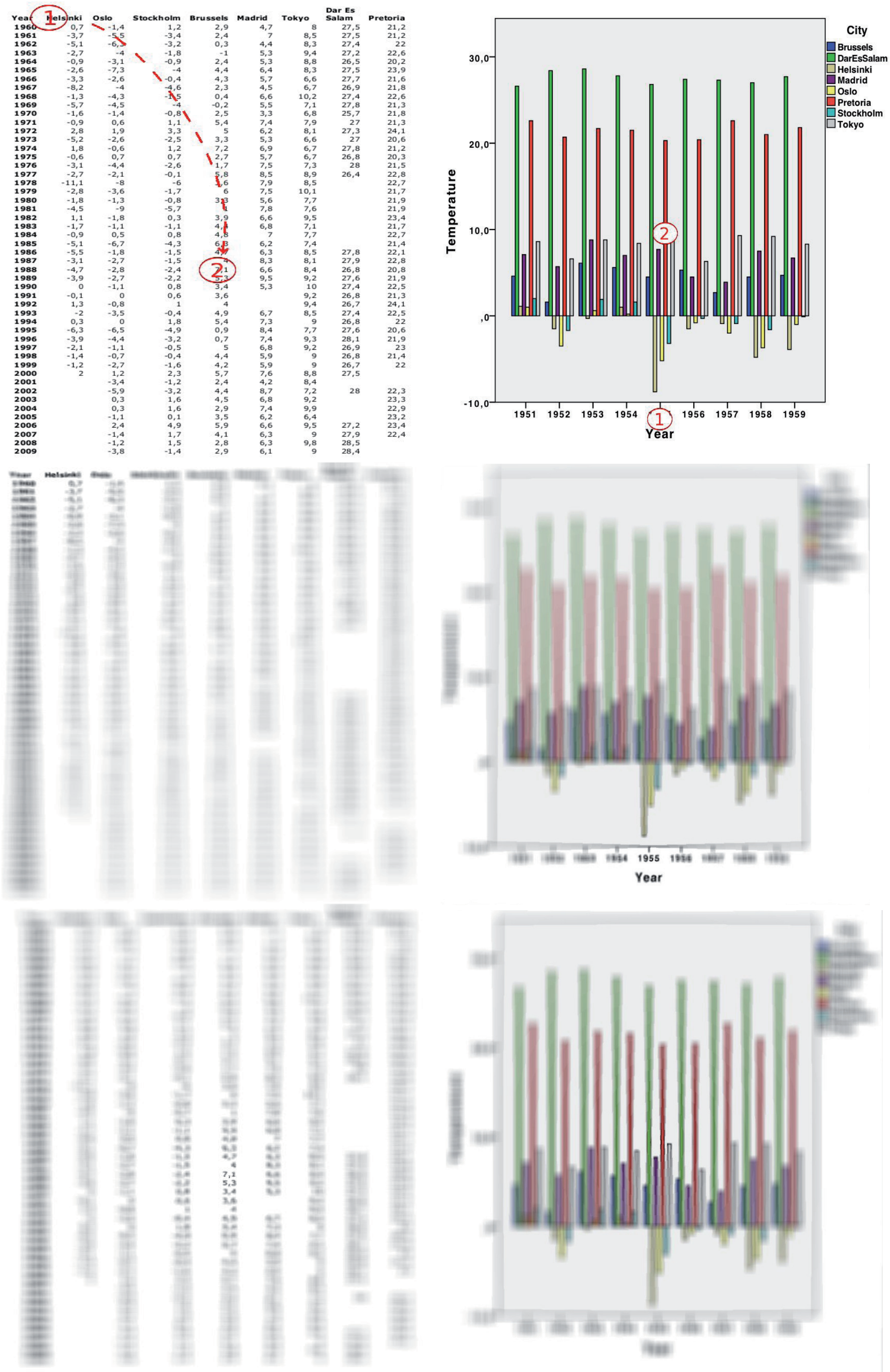
The use of visualizations is influenced by very many subjective (e.g., personality, experience, skills) and contextual (e.g., motivation, fatique, expectations, culture) factors, but the objective of making general predictions must compromise many of these factors. Fortunately, situations and user’s have much in common and users' abilities to understand visualized data are limited by similar perceptual ['vision'] and cognitive ['CS'] processes. The discussed perceptual limitations are often related to *eye saccades,* the rapid shifts of visual gaze point. Due to the limits of visual acuity, users of data intensive visualizations pay attention to spatially distributed details one at a time (Fig. 2). The quality of perceptually extracted visual features is different at peripheral vision compared to that at central vision, and this difference is specific to visual feature (color, line orientation etc.) evaluated. The narrow extent of high-acuity vision is a surprise to many people as it is approximately the width of a thumb at normal viewing distance from the computer monitor (O'Shea 1991). In Figure 2, the blurred peripheral vision provides outlines of data values with the bar graph (analogical), but not with the numeric table (digital). Thus, peripherally presented visual information can determine the search performance. In these examples, the perception of even static figures can be sequential (*perceptual dynamics*). Another cause for the sequentiality in interactive visualization is the *interface dynamics* ['Viz'], as only parts of the data are presented at a time according to manual input ['operation']. The term interaction in this thesis includes both of these factors, perceptual and interface dynamics, although user-interaction models typically discuss only the latter.

The limits of visual acuity and sequentiality are the reasons why the following three subtasks are so influential for the usability of visualizations. First, peripheral vision is utilized for creating an outline of different parts of the visualizations (Sect. 5.1). Second, the visual outline is used in attempts to locate and focus on the relevant details with high-acuity central vision (Sect. 5.1 - 5.4, 6.2). Third, visual memory is used to integrate over these "snap shots", so that the details can be related to one another (Sect. 5.1, 6.1). These perceptual subtasks are important when user attempts to understand data-intensive visualizations because everything cannot be seen at once. The first subtask was used as design principle in the election services (Publ. I & II) and the last two are discussed in the following Chapters as evaluation criterion for graphs in psychophysical experiments (Publ. V & VI).

The integration task is difficult because it is not easy to memorizing the detailed information across gaze shifts (Publ. IV). Human working memory capacity is limited to as few as four unrelated items (Cowan 2000; Vogel et al. 2001, unlike the legendary 7 in Miller 1956) and the conceptual challenge ['CS'] is to relate the details. The human abilities to relate details of multi-dimensional data are analyzed with respect to Conceptual Spaces (CS) theory (Gärdenfors 2000). Thus, human understanding is analyzed only with respect to the aspects most immediately related to the use of visualizations, which excludes for instance the long-term development from childhood.

According to CS theory, human concept formation and categorization takes the perceived multi-dimensional information as input for subjectively constructed low-dimensional mental representations. CS theory generalizes the knowledge about human categorization of simple perceptual stimuli to very abstract domains such as politics. In both cases, presented information takes form of high dimensional data points, the data point is mapped to low dimensional conceptual space, and categorized according to the distances to each of the alternative categories. This thesis evaluates the empirical evidence for the three fundamental principles of CS theory. First, conceptual information is mentally represented in a low-dimensional space. Second, the mental representation consists of newly constructed dimensions instead of the perceptual input dimensions, if the input dimensions are numerous. Third, categorizing an item to a certain subpart of the mental space explains the decisions based on multi-dimensional information. The categorization functions by estimating which is the best-match between the item to be categorized and central tendencies (prototypes) of items in different subparts of mental representation. The implications of the theory are analyzed for instance when the users of political applications (Ch. 2) try to understand opinions of hundreds of candidates on different topics, so that they can choose the best candidate in the elections. For instance, voter might position candidates on a subjectively constructed conceptual dimension of 'environmental friendliness' based on several political issues. The constructed dimensions might not be directly related to any single dimension of input data, and they can be context dependent. Similarly, the scale for the conceptual 'hotness' was argued to be different for the tap water and the bath water (Gärdenfors, 2000, p.120).

The challenge with the multi-dimensionality of data is that a person with limited cognitive capacity might adopt suboptimal strategies without acknowledging it. Instead of construction low dimensional representations, one might simply select a subset of the presented attributes. For instance, the consequence of processing only some of these attributes often lead to stereotypical and extreme evaluations (e.g., of food and people by Linville 1982). To avoid these consequences, all the attributes need to be processed and it helps if they can be represented in computationally more-structured manner either by the human mind ['CS'] or by the computer ['DR' in Fig. 1].



Year	Helsinki	Oslo	Stockholm	Brussels	Madrid	Tokyo	Dar Es Salam	Pretoria
1960	0,7	-1,4	1,2	2,9	4,7	8	27,5	21,2
1961	-3,7	-5,5	-3,4	2,4	7	8,5	27,5	21,2
1962	-5,1	-6,3	-3,2	0,3	4,4	8,3	27,4	22
1963	-2,7	-4	-1,8	-1	5,3	9,4	27,2	22,6
1964	-0,9	-3,1	-0,9	2,4	5,3	8,8	26,5	20,2
1965	-2,6	-7,3	-4	4,4	6,4	8,3	27,5	23,9
1966	-3,3	-2,6	-0,4	4,3	5,7	6,6	27,7	21,6
1967	-8,2	-4	-4,6	2,3	4,5	6,7	26,9	21,8
1968	-1,3	-4,3	-1,5	0,4	6,6	10,2	27,4	22,6
1969	-5,7	-4,5	-4	-0,2	5,5	7,1	27,8	21,3
1970	-1,6	-1,4	-0,8	2,5	3,3	6,8	25,7	21,8
1971	-0,9	0,6	1,1	5,4	7,4	7,9	27	21,3
1972	2,8	1,9	3,3	5	6,2	8,1	27,3	24,1
1973	-5,2	-2,6	-2,5	3,3	5,3	6,6	27	20,6
1974	1,8	-0,6	1,2	7,2	6,9	6,7	27,8	21,2
1975	-0,6	0,7	0,7	2,7	5,7	6,7	26,8	20,3
1976	-3,1	-4,4	-2,6	1,7	7,5	7,3	28	21,5
1977	-2,7	-2,1	-0,1	5,8	8,5	8,9	26,4	22,8
1978	-11,1	-8	-6	1,6	7,9	8,5		22,7
1979	-2,8	-3,6	-1,7	6	7,5	10,1		21,7
1980	-1,8	-1,3	-0,8	3,3	5,6	7,7		21,9
1981	-4,5	-9	-5,7	4	7,8	7,6		21,9
1982	1,1	-1,8	0,3	3,9	6,6	9,5		23,4
1983	-1,7	-1,1	-1,1	4,4	6,8	7,1		21,7
1984	-0,9	0,5	0,8	4,8	7	7,7		22,7
1985	-5,1	-6,7	-4,3	6,3	6,2	7,4		21,4
1986	-5,5	-1,8	-1,5	4,7	6,3	8,5	27,8	22,1
1987	-3,1	-2,7	-1,5	4	8,3	8,1	27,9	22,8
1988	-4,7	-2,8	-2,4	7,1	6,6	8,4	26,8	20,8
1989	-3,9	-2,7	-2,2	5,3	9,5	9,2	27,6	21,9
1990	0	-1,1	0,8	3,4	5,3	10	27,4	22,5
1991	-0,1	0	0,6	3,6		9,2	26,8	21,3
1992	1,3	-0,8	1	4		9,4	26,7	24,1
1993	-2	-3,5	-0,4	4,9	6,7	8,5	27,4	22,5
1994	0,3	0	1,8	5,4	7,3	9	26,8	22
1995	-6,3	-6,5	-4,9	0,9	8,4	7,7	27,6	20,6
1996	-3,9	-4,4	-3,2	0,7	7,4	9,3	28,1	21,9
1997	-2,1	-1,1	-0,5	5	6,8	9,2	26,9	23
1998	-1,4	-0,7	-0,4	4,4	5,9	9	26,8	21,4
1999	-1,2	-2,7	-1,6	4,2	5,9	9	26,7	22
2000	2	1,2	2,3	5,7	7,6	8,8	27,5	
2001		-3,4	-1,2	2,4	4,2	8,4		
2002		-5,9	-3,2	4,4	8,7	7,2	28	22,3
2003		0,3	1,6	4,5	6,8	9,2		23,3
2004		0,3	1,6	2,9	7,4	9,9		22,9
2005		-1,1	0,1	3,5	6,2	6,4		23,2
2006		2,4	4,9	5,9	6,6	9,5	27,2	23,4
2007		-1,4	1,7	4,1	6,3	9	27,9	22,4
2008		-1,2	1,5	2,8	6,3	9,8	28,5	
2009		-3,8	-1,4	2,9	6,1	9	28,4	

Fig. 2. Tabular and bar graph representations for human visual system. The three images on the left represent differently filtered versions of a table. The eye fixations (numbered circles) and a saccade (dashed line) are marked in red on the figures at the top row. Below the top row table, the two images are filtered according to the fixations[1]. The first fixation (at the middle) is directed to a label and the second (at the bottom) to an illustration of a data value. These examples illustrate the importance of peripheral vision, as it provides an outline of quantities in the case of bars.

Overcoming the challenges within the framework

Visualizations are an important way to overcome mental limitations. Figure 2 shows an example of the ways in which visualized forms can be more adapted to the features of human visual system than alphanumeric forms. Most importantly, diagrammatic form preserves geometric relations that are only implicit in sentential symbolic forms (Larkin & Simon 1987). These relations are by themselves more suitable for humans, if the Conceptual Spaces theory (discussed in next Section) is valid. By contrast, the argument is less compelling if the internal representations would be formed as propositions (logical statements) as was suggested by Pylyshyn (1986) and if there is nothing special about visual modality. In an alphanumeric representation, one often needs to re-read the text to find a detail, but in diagrammatic representation location can index data structures (Larkin & Simon 1987). This idea is evaluated in this thesis using *visual search,* an experimental psychology procedure. Visualizations can *externalize* information processing from the required symbolic and logical inferences to perceptual ones (Larkin & Simon 1987; Koedinger & Anderson 1990; Kirsch 1995; Cheng 1999; Scaife & Rogers 1996) as the visualization performs the processing through steps of Figure 1. The general idea of externalization is that information processing is done by other means than user's cognition.

In addition, visualization includes the decision of how to do the visual representing. An influential statistician, John Tukey, and the founder of cognitive psychology, Ulrich Neisser have both endorsed similar views in 1970s according to which complex issues are understood by interactive explorative activity (Tukey 1977; Neisser 1976). According to Tukey, symbolic representations are efficient for small number of data, but understanding of large quantities benefit from creating hypotheses from an overview of the data and confirming them from the details. This thesis concludes that periphery of visual field is useful for presenting such overviews, and high-acuity central vision is useful for the confirmations. Neisser's view is

[1] Gaussian blur was applied and its size changed according to visual acuity (Vernier acuity in Duncan & Boynton 2003) at different distances from gaze point. The same procedure was applied to bar graphs on the right.

consistent with this as it emphasizes the iterative nature in confirming the expectations about the environment.

A likely reason why peripheral vision has not been emphasized enough in visualization is that the related processes are unconscious. People are not commonly aware of their gaze point and cannot self-reflect about these processes. Visual acuity, for one, is very much limited in visual periphery but people typically have an impression that they would not have such limitations (O'Regan 1992; Publ. IV). The reason is that 'just-in-time' eye movements compensate for this limitation (Johansson et al. 2001; Hayhoe et al. 2003; Land 2004; Henderson & Castelhano 2005) creating so called *virtual representation* (Rensink 2000). The idea has been illustrated by a metaphor of Internet, in which the content might seem to be in your possession, when in fact it is retrieved instantaneously from servers. The thesis evaluates visualizations as such virtual representations that extend human memory and processing capacities.

1.2 Research questions

Q1. What kind of interaction with visualizations would support understanding of multi-dimensional information?

This is the most general question of the thesis for which the interaction framework (Fig. 1) is introduced. The framework was derived from experiences with a practical application (Ch. 2 & Publ. II) and theoretical generalizations from the cited studies about human vision, human decision-making, neuroscience, and use of visualizations. The understanding of visualizations depends on interaction of very many contextual, perceptual, cognitive, motivational, cultural aspects. The interaction framework identifies generalities by focusing on perception and cognition independent of context and subjective aspects the user.

Based on the framework, three new and general princples concerning visualizations were presented (Ch. 3 & 7). The more specific contributions to answering this research question are discussed in terms of the other research questions. The first general way is to rely on unconscious and intuitive processing, as opposed to logical reasoning. The utility of intuitive processing is evaluated in several decision-making situations, for instance with shopping (Dijksterhuis et al. 2006) or value judgment (Levine et al. 1996).

The second way is observing consequences of the actions that alter the visual perceptual input from the visualization. User can press a key and change some parameters controlling the way the data is drawn on computer screen or simply look at a different location on the screen. As a consequence the visual detail of interest can change or remain the same. To support such approach for understanding the data, it is argued that this is a fundamental way in which humans understand the external world. This idea is supported by a review of visual neuroscience that shows that a large portion of visual system is dedicated to such inferences. As a consequence, the importance of learning visual interaction skills is emphasized in addition to the more commonly acknowledged objective of obtaining knowledge from the represented data. Perhaps, gaze shifts have generally not been considered as means for the user to interact with the information flow from the visualizations. Even so, gaze shifts are here included as an interaction skill because they importantly influence user's perception and conceptualization of the visualized data. Looking at irrelevant locations can delay or prevent understanding of large data.

The third way a designer of visualizations can ensure that users focus on the most relevant details by relieving the users from as many cognitive processes as possible. This is called *externalization* as the tools are performing operations that would otherwise be left for the users' minds. A user of visualization is conceptualized as a part of the larger information processing system (Liu et al. 2008). This thesis goes in much deeper into this topic than the earlier articles about externalization with visualizations.

The above evidence for the proposed framework of interacting with visualization is in many ways limited due to the breadth of the issues involved. The evidence is indirect (from cognitive neuroscience), theoretical (externalization and intuition), and sample only small set of cognitive processes (analysis of intuitive decisions). However, the framework also provides very reliable conclusions about peripheral vision founded on direct and converging evidence from different fields of research (physiology, psychophysics, neuroscience, and other behavioral experiments). This aspect is later utilized in a more detailed visualization method (see Q4).

Q2. Can mental processes alone simplify multi-dimensional perceptual input from external world analogously to dimension-reduction algorithms?

The human concept formation abilities with respect to this question are evaluated by reviewing literature on empirical findings with natural data sets (colors, visual object categories, olfaction, touch, and taste) in Chapter 4. This complements the supportive evidence reviewed for Conceptual Spaces theory (Gärdenfors, 2000).

The evidence includes findings that humans categorize high-dimensional sensory information according to few emergent dimensions. For instance, the number of dimensions used to differentiate scents is much fewer than that of odor receptor types (~ 1000) (Buck & Axel 1991). However, this finding would not suffice as support if the person arbitrarily selected some of the sensory dimensions. Simplification in this research questions refers to the ability to construct low dimensional representations based on high dimensional input. Simplification is an abstraction process that might select and weight the input dimensions, or it might create entirely new emergent dimensions based on the input dimensions.

Q3. What kinds of mental representations are thought to be central to human concept formation and the commonly used dimension-reduction algorithms?

Main purpose in Chapter 4 is to apply Conceptual Spaces theory (Gärdenfors, 2000) for selecting what to represent in visualizations. The underlying idea is that if the humans are naturally thought to utilize certain forms of information representations, they are more likely to be comprehended easily compared to any other alternatives. First, the theory is recognized to be appropriate for the purpose. To my knowledge, it is the only available theory of human concept formation that focused on multi-dimensionality and is also explicit enough for implementations (e.g., unlike Mental Models theory of Johnson-Laird's 1980; 1981, EPIC architecture of Kieras and Mayers, 1997, or ACT-R theory of Anderson et al., 2004). Second, the aspects relevant for multi-dimensional visualization are identified and evaluated. For instance many linguistic aspects were excluded. The evaluations included constructing a process model based on the theory (Publ. III) and comparing the properties of the theory to those of dimension-reduction algorithms. Third, the relevant aspects are analyzed for purposes of applications and especially for the political one (Publ II). As an answer to the question, three properties were identified: metric spaces, dimension reduction, and prototypes.

Q4. What kind of visual representation of the preprocessed data would guide user's eyes and attention to the most relevant locations?

The research question stems from the proposed sequential framework (Fig. 1) and is answered by reviewing vision research, first directly (Sect. 5.1) and then with respect to visual searching (Sect. 5.2 – 5.4). First, the review supports a conclusion for visualization that peripheral vision is a useful information channel for the guidance of focused attention and gaze shifts. Second, visual search results are used for explicating visual design features that determine the possibilities for such guidance.

A generally accepted visualization technique is to represent detailed information with context or overview so that user understands which details are in question (Furnas 1981; 1986; Shneiderman 1996). Accordingly, this thesis suggests a method of representing the core structures from dimension reduction as an overview in peripheral vision. This method serves the purpose of externalization (see Q1 above) in two ways. The first is to relieve user from some aspects of constructing multi-dimensional representation in mind by creating simplifications. The second way is to relieve user from memorizing the details by designing them to be quickly and effortlessly accessed by gaze shifts. These externalization aspects were included in the design of application that is described in Chapter 2, criteria for rapid visual access are outlined in Chapter 5, and Chapter 6 shows how the rapid access criterion can be evaluated with psychophysical experiments.

Q5. What are the "bottlenecks" of perceiving visualizations of multi-dimensional information?

In another words, each visualization type has pros and cons, and the objective of understanding multi-dimensional visualization is used to determine what perceptual and cognitive tasks determine efficacy of visualizations. Two sources of information were used to identify the critical tasks, the proposed framework of interaction (also Publ. IV) and behavioral studies about the use of visualizations (discussed in Publ. V & VI). As a result, searhing and integrating tasks were identified as critical. These were proposed instead of the traditional criteria, perceptibility of absolute values or ratios (Cleveland & McGill 1984; 1985). The critical tasks were more likely to be difficult, frequent, and influenced the time needed for gathering the required information from visualizations than the traditional criteria.

For the search criterion, the relevant details need to be located quickly from the large quantities of data. The high numbers of data values often result in high number of visual elements in the display. This might lead to compromised usability if attending to relevant details instead of irrelevant ones becomes more difficult. For the integration criterion, the data values on different dimensions need to be integrated into meaningful summarizing concepts or chunks of representations. For instance, users of the example application (Ch. 2) typically attempt to summarize and compare the opinions of different political candidates with respect to different issues. Similarly customers, of products with multiple features, somehow have to represent and compare many different qualities of individual alternatives. The importance of comparing pairs of items with respect to many features became apparent during the qualitative research of the example application. The integration subtask determines the ease of accounting all the available features. The outcome of answering this research question is formulating criteria for evaluation of visualizations. The more detailed discussion of the criteria can be found from Publications V and VI.

Q6. How should multi-dimensional visualizations be evaluated quantitatively in psychophysical experiments?

The thesis develops methodology to evaluate graphs based on methods of vision research. Both the framework and the review of visual search experiments (Ch. 5) were used to develop methods capable of measuring different levels of efficiency of test subjects' performance associated to the utilized graph type (Publ V & VI).

The different objectives lead to similarities and differences. Vision research evaluates the functioning of perceptual processes, and Publications V and VI evaluate the effectiveness of graphs in a given task. The above-mentioned subtasks, locating and integrating, can be operationalized into experiments that quantitatively compare different graphs (Chen & Yu 2000). *Operationalization* here means that important aspects of real visualization use are modeled by simpler controlled experiments. The operationalization process is not immediately obvious. First, most studies of visualizations are qualitative (e.g., Perrer & Shneiderman 2009, but see Chen & Yu 2000). Second, the contributions of the perceptual characteristics in complex displays are difficult to evaluate (see House et al. 2006 for some solutions). Third, studies with simple graphs have evaluated very different subtasks than the ones proposed here (see Publ. V for review). The proposed subtasks are argued to be *ecologically representative*, as they model the real use of visualizations and challenges to understand multi-dimensional data.

The above presented research questions progress from the development of empirically supported framework for conceptualizing the interactive use and understanding of visualizations toward the more practical evaluations of visualizations. The main contribution of this thesis is to explain the challenges of human perception and understanding with the framework, identify criteria to evaluate those challenges, and develop methodology to operationalize the criteria in controlled experiments. Without such efforts, the user experiments might evaluate irrelevant aspects of visualizations. Table 1 outlines how different Publications contribute to the research questions.

"Object displays for identifying multidimensional outlines from crowded visual periphery"

"Integrating complex information with object displays: Psychophysical evaluation of outlines"

"Eye movements and the delusion of high acuity visual field"

"Category learning by formation of regions in conceptual spaces"

"Exploring political agendas with advanced visualization and interface tools"

"Enahancing usability of the similarity map for more accessible politics"

		Pub-1	Pub-2	Pub-3	Pub-4	Pub-5	Pub-6
What kind of interaction with visualizations would support understanding of multi-dimensional information?	Q1	x	x		x	x	x
Can mental processes alone simplify multi-dimensional perceptual input from external world analogous to dimension-reduction algorithms?	Q2			x			
What kind of mental representations are thought to be central to human concept formation and the commonly used dimension-reduction algorithms?	Q3			x			
What kind of visual representation of the preprocessed data would guide user's eyes and attention to the most relevant locations?	Q4				x		x
What are the "bottlenecks" of perceiving visualizations of multi-dimensional information?	Q5				x	x	x
How should multi-dimensional visualizations be evaluated quantitatively in psychophysical experiments?	Q6					x	x

Table 1. Relations between different research questions(Q1 - Q6) and Publications (Pub-1 - Pub-6).

1.3 Structure of thesis

The structure of the thesis is as follows. Chapter 2 describes the design of an example application of multi-dimensional visualization in the field of politics. The most relevant aspect of the application is the use of dimension-reduction algorithms in services targeted for general audience. This application serves as a reference throughout the following discussion. Chapters 3 and 4 discuss the proposed principles for interactive multi-dimensional visualization. Chapter 3 discusses the interaction related conceptions and the related behavioral studies. Chapter 4 discusses Conceptual Spaces theory and the spatially represented results of dimensions reduction. Chapter 5 explicates visual design features that influence peripheral vision and visual searching performance. In general, two critical perceptual and cognitive subtasks, searching and integrating, are identified as evaluation criteria for graphs. Thus, visual search experiments are also discussed as an evaluation method. Chapter 6 discusses the evaluation of diagrams based on the novel subtasks. The relationship between Chapters, Publications, and later-proposed principles is presented in Table 2. The proposed principles and other conclusions are discussed more thoroughly in the final Chapter of this summary.

Table 2 Main contributions in chapters

Ch.	Publications	Research questions	Proposed principles	Other contributions and conclusions
2	[I], [II]	Q1	P3	feasibility of P3 in a full-scale example, benefits of sector diagrams for DR
3	[IV], [V], [VI]	Q1, Q5	P1, P2	role of unconscious processing, externalization with gaze shits
4	[III]	Q2,Q3	P3	emprirical evidence for human DR a process model to support CS theory CS motivated DR properties for visualization
5	[IV,VI]	Q4, Q6, (Q5)	P3	identify critical visual features for locating, gaze point control for evaluation methodology
6	[V,VI]	Q5,Q6	-	sector diagrams suitable for externalization the evaluation criteria can be operationalized

Principles from Section 7.1:
P1 = Visual representation can be designed to relieve user from otherwise necessary cognitive load.
P2 = Visualization adapted to human vision serve both mental representations and actions.
P3 = Simplifications of dimension-reduction (**DR**) algorithms can guide user by providing outlines of data through peripheral vision.
for research questions see Section 1.2
CS = Conceptual Spaces

2 A CASE STUDY OF INTERACTIVE MULTI-DIMENSIONAL VISUALIZATION

Overview of Chapter

This chapter presents a case of interactive multidimensional visualization, which is discussed throughout the summary. The case example illustrates theoretical conceptions and shows their relevance. In general, the discussed objectives of conceptual simplicity and ease of perceptual inferences are likely to be more important when non-expert users, instead of scientists, are trying to understand complex visualization. This case study is exceptional because it involves a service targeted for common people while utilizing dimension reduction algorithms. The methods for developing the service are described in Publications I and II (also in Kerminen et al. 2000; Kaipainen et al. 2001; Berg 2004). Section 2.1 outlines the theoretical objectives of the service and Section 2.2 describes the resulting service.

2.1 Purpose of the designed election engine application

The main objective of this thesis is to provide visual information in a more understandable form. One of such ways is to use dimension-reduction (DR) algorithms. Computer science experts have been the standard user-group of DR algorithms (such as MDS and SOM). A central argument of this thesis is that inexperienced users can also benefit from these algorithms. This chapter provides an exceptional example when non-expert users interactively study this kind of visualizations. Thus, the use of that visualization service can also be considered as an interesting object of an ethnographic study (e.g., Shneiderman & Plaisant 2006; North, 2006; Perer & Shneiderman 2009). The service was especially designed to alleviate the usability difficulties resulting from the lack of experience with multi-dimensional information in this user-group.

The reason to use these computationally more complicated systems is that treating this information as low dimensional will lead to decisions of poor quality. They might selectively or randomly attend only to a subset of task relevant information. In the analogous example from shopping (Dijksterhuis et al. 2006), it would be detrimental to simply ignore, for instance, some of the dimensions of the car alternatives. The application context of Publications

I and II was elections. Let us speculate a little about the consequences of cognitive limitations with the political data. The users with such limitations would be choosing their candidates based on four simplified issues pre-selected by political journalists. This conclusion would be based on the assumption that people cannot typically hold in mind more than four unrelated items (Cowan 2000 for review of memory research) if they are not motivated for laborious task of relating (termed *chunking*) the items. Four issues can hardly represent the complexity of issues involved in representative democracy. This was taken to be the challenge of the discussed application that involved dimension reduction. The application attempted to increase users' ability to include different aspects or perspectives of the political issues into consideration when making the complex decisions about voting.

2.2 Short description of the published service

The election related services consisted of four main components (Fig. 3) designed for exploration of candidate data (Publ. I & II). The two traditional election service components included were (i) an interface for answering 20 or 25 five-alternative-choice questions (on left) and (ii) list of the best-matching candidates (on right). The novel visualizations components included were (iii) a similarity map based on SOM-algorithm (in the middle) and (iv) two sector diagrams (below the map). The similarity map spatially organized all the candidates for browsing. The candidate data consisted of the decisions the candidates would make with respect to the same questions shown to the user. The candidates also had an opportunity to provide verbal justification for their decisions.

A mentioned theoretical problem with the application is that users did not know what SOM is, how it operates, or even that it was used. However, the behavior of at least some of the test users indicated that they understood the principles of preprocessed data provided by SOM algorithm. Some realized that more similar candidates are closer to one another (as in topographies of Conceptual Spaces discussed in Sect. 4.2). This realization had profound effect on their exploration. For instance, some users decided to look for candidates of other extreme to see the whole spectrum. The coloring reflected the spectrum of similarities. Dark green represented similar opinions with the user and red was the opposite. The coloring reflected only the questions answered by user, and user could simply skip irrelevant questions. The default location was the most similar candidates to user's own opinions (as in Fig. 3). Thus, even if one does not realize the underlying organizational principle, it is likely to be more informative to browse these candidates instead of random candidates.

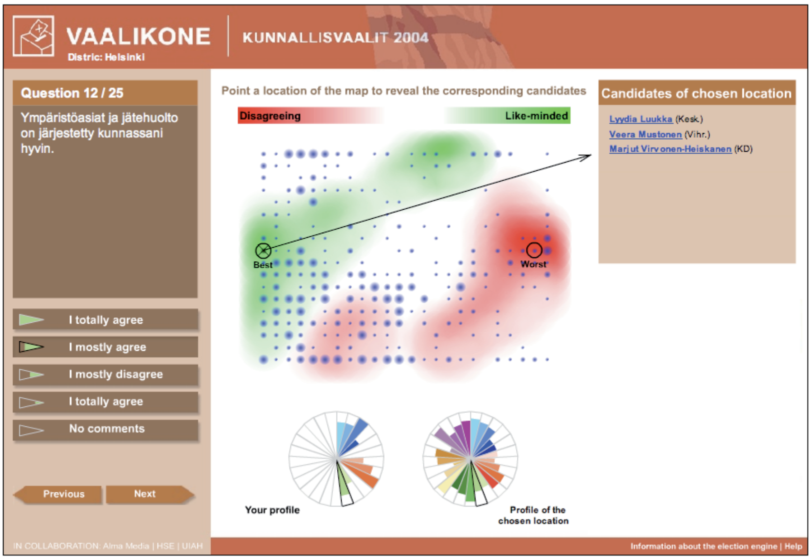


Fig.3 A snap shot from published application.

From the two sector diagrams, one represented opinions of the user and the other represented those of best-matching candidate (or other selected candidate). Sectors represented different political issues (dimensions) ordered according to larger themes. The length of a sector represented the agreement with the corresponding question, which was formulated as a statement.

Data with large number of candidates is problematic for visualization because everything cannot be conceptualized at once. In the discussed application, user had to choose the most interesting or relevant details. This follows the details-on-demand principle (e.g., Shneiderman 1994; 1996), according to which everything is not shown at once and the details are only revealed when the user selects them (see Sect. 3.1). This is consistent with the memory offloading principle (Sect. 3.3) because environment holds the information until user needs it. User could browse the following hierarchy: (i) the map represented the whole data, (ii) a location on the map represented few similar candidates, and (iii) the choices of a selected candidate were supplemented by written justifications.

The decision to include the sector diagrams had only theoretical grounds when the service was published (Publ. II). The decision was based on basic research findings about visual perception of low-level visual features, which are discussed in Chapter 6. Our hypothesis was that sector diagrams would be easy to locate and compare because of their visual outline. The later experiments (Ch. 6) applied the methods from a basic research of visual search to evaluate the hypothesis (Publ. VI) and same time developed new evaluation criteria for the task of comparing individual graphs (Publ. V & VI).

2.3 Dimension-reduction for novice users

Overview of Section

This section discusses the ways in which the application followed the later suggested information design principles motivated by Conceptual Spaces theory (see Ch. 4). In practice, the general visualization principle of simplicity (Tufte 1983) is applied to the understanding of multi-dimensional data. For the discussed application, this means helping non-expert users to overcome the difficulties with large data quantities and multi-dimensionality. The offered solutions for the challenges are:

(i) converting assessments of algorithms into evaluations of visual shapes, and

(ii) using details-on-demand principle to reduce concurrently presented information.

These solutions were the special methods, and they implement the more general ideas of interactive visualization and externalization discussed in Chapter 3.

The applications utilized a dimension reduction algorithm (SOM) that Gärdenfors claimed to function analogously to human conceptualization processes (see Ch. 4). Visualized output of the dimension reduction algorithms can provide perceptual input for the users' mental representations. The core characteristics were also emphasized in the way the output was visualized. For instance, the gradient coloring of the map reflected the continuity of the spatial order, and the increased saliency emphasized the opposite ends of the scale.

Simplification is helpful if non-experienced users want to understand large quantities of multidimensional information. According to Edward Tufte (1983) unnecessary details should be reduced. In this thesis, the simplicity is motivated by the capacity limits of perceptual and cognitive processes to ensure the quality of decisions. The simplicity discussed in this section involves only the dimensionality of data. In the multi-dimensional visualization input data is high dimensional, and it is simplified by reducing the dimensionality algorithmically.

Then, how does the user group influence the requirements of visualization? Scientists and statisticians are able to evaluate the algorithms for a given data, unlike the typical voters. Yet, the voters need to be able to understand and control multi-dimensional decision aids, or else they avoid using them (Shneiderman 2002). There are good reasons for the voters not to trust the algorithms, even if the algorithms are not misused. By their nature, multi-dimensional problems do not have unique solution that could be unequivocally trusted. Different dimensions give support for a different set of decision alternatives and some of these influences need to be compromised. Therefore, the compromise may need to be represented explicitly (in visual form) so that user can acknowledge and approve it. User might not approve the compromise and might want to see consequences of changing priorities of different dimensions. The general consequence has been that users from very different professions (even engineers) have not trusted on such decision aids because they have not been able to evaluate the underlying assumptions (Kleinmunz 1990; Laine 2003). These reasons could explain why experts have been the only real user-group for multi-dimensional visualizations of the past.

The presented resolution to the problem in Publication II was to include interaction with simple graphs. Simple graph present undistorted details and does not reduce the original dimensionality. For instance, original data is used as an independent source to evaluate the similarity clusters on spatial representation of the SOM visualization. If SOM is not useful for the given data, the items (here: candidates) within a cluster would not share values for individual dimensions (here: degree of agreement to statements). All the DR algorithms at least assume that the reduction to only few dimensions is meaningful. It is certainly possible that the few underlying dimensions capture only small part of the whole variance in the data.

In terms of visualization, 'validity assessments' are converted to assessments of visual similarity (Fig. 4.). The dimension reduction is valid if user has an intuition that the simplified (dimensions reduced) information represents the reality (original data) in a meaningful way. The data in Figure 4 is from SOM-based election service for Finnish EU elections 2004 (Publ. I & II). In the service, user expressed opinions to the same questions than candidates and these opinions were compared. The users did this comparison by evaluating overall shapes of sector diagrams. Sector diagrams do not represent any additional computational operation but represents the raw data of individual entities (here: candidates) as simple graphs. These graphs are later (Sect. 6.1) termed object displays due to their perceptual characteristics, and they are used to facilitate attending to the most interesting details.

A novel approach of this thesis is to use such object display diagrams to clarify the consequences of dimension reduction. The simple graphs can help to reveal the underlying data related to later discussed prototype structure and categories. In that case, part of the demanding cognitive process of constructing conceptual representations is externalized to decision aid service.

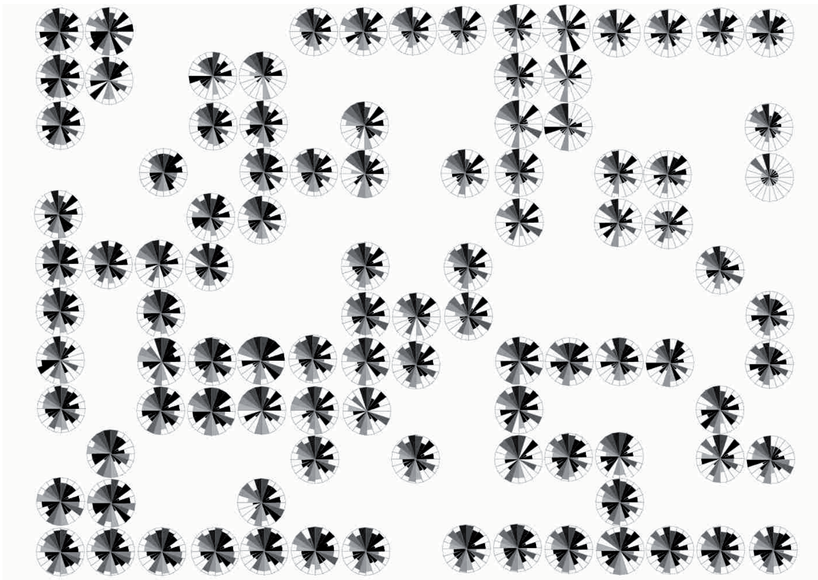


Fig. 4. A set of sector diagrams. Sector diagrams represent the data before dimension reduction and they are organized according to SOM. In real application, the sector diagrams would be visible to the user one at a time when selected by a cursor on top of the map (not shown here). The adjacent diagrams have more similar shape than the diagrams at the opposite end. This is only true if SOM has validity with the give data. For instance, the sectors are generally longer at bottom left as the corresponding candidates have agreed more with the represented 20 statements. Abilities to compare and locate interesting diagrams, are required. They are evaluated in psychophysical experiments of Chapter 5. This figure does not have colors because the performance in those experiments was improved when the colors were removed.

2.4 Summary of contributions

The most important contribution of the work described in this chapter is the dimension-reduction-based service utilized by wide audience (Publ. I & II). The two most important occasions when this kind of service was made available were before the Finnish EU elections in June 2004 and before the Communal elections in October 2004. On the second occasion, the service was used close to 100.000 times (see Publ. II). The popularity is important because the researchers of advanced visualizations have been greatly concerned about impact of their work (van Wijk 2006). The larger user-groups have typically not been eager to adopt complicated dimension-reduction tools offered in Internet. Therefore, much can be learned from this successful application type.

This chapter has explained details of a case study of service. The novel objective of this service has been to deepen users' understanding influencing their subjective selection criteria for political candidates. The use of a dimension reduction algorithm (SOM) facilitated understanding of multi-dimensional data. Special visualization means were developed so that novice users could visually evaluate the validity of the algorithmic simplification of the data. Consistency in shapes of sector diagrams with spatial location of SOM was an indicator of validity (Fig. 3 & 4). The thesis presented a general and novel idea for helping inexperienced users to trust on the dimension reduction output by adding object display graphs next to the output (Ch. 6). The idea is to convert some mathematical validity assessments of the algorithms into assessments of visual similarities. Furthermore, the large quantity of data was handled by presenting only those details that user selected from a hierarchically organized structure (all candidates / a subset of similar candidates / opinions of one candidate). The following Chapters 3 and 4 address the theoretical framework of interactive visualization, which motivated the design decision.

3 INTERACTION AND DATA-BASED INTUITIONS TO OVERCOME COGNITIVE LIMITATIONS

Overview of Chapter

This chapter answers the first and most general research question of Section 1.2:

What kind of interaction with visualizations would support understanding of multi-dimensional information?

This is done by discussing the visualization principles related to unconscious learning, interactive use of visualizations, and external cognition. These principles circumvent the typical cognitive limitations related to understanding of multi-dimensional data, most notably the use of working memory to integrate large quantities of details into coherent representations.

3.1 Intuitions from interactive exploration

Overview of Section

The first section of this chapter (3.1) discusses the first two of the mentioned principles. Human intuition is too wide topic to be discussed here thoroughly. Here, intuition refers to unconscious understanding of the visualized data. Unconscious learning refers to the opportunity to increase understanding of the data by interacting with it. The suggested theoretical view is that unconscious learning is significant factor in interactive visualization because users are often not consciously aware of the information they gain from visualizations. This section provides tentative empirical support for the notion that humans make often better multi-dimensional decisions unconsciously than consciously reasoning.

The benefit of providing external medium for cognitive processing is that user's efforts are relieved for constructing mental representations of the data. The medium is interactively accessed through the perceptual system. This section claims that understanding of multi-dimensional visualizations can be facilitated by increasing (i) data-driven intuition (unconscious cognition) and (i) structure revealing interaction. Data-driven intuition is grounded on the visualized information, as opposed to being unjustified or arbitrary. Interaction is an important way to do this grounding, but also static images can facilitate intuition supported by visual perceptions. The importance of interaction is also supported by Section 3.2, which shows that important part of human visual cognition is interactive. Unconsciousness is also another aspect characteristic for human vision. Even historically, the idea of unconscious cognition originated from research of low-level visual perception. The idea came from Hermann von Helmholz's theory of visual inferences in 19th century, instead of later theories of self by Sigmund Freud (1923).

Intuition grounded decision have been found more appropriate for multi-dimensional data than consciously deliberated decisions. This conclusion has been supported by following studies of shopping, face recognition, value judgments, evaluation of taste, and planning of college courses. To begin with, conscious deliberation was found to be useful for only low-dimensional shopping decisions (Dijksterhuis et al. 2006). In contrast, intuition led to better decisions, when subjects had to buy items with very many potentially important features. However, later more controlled replications (Newell et al. 2009) of these studies contest some of the original claims. In these experiments, long lasting periods of non-attentive contemplation, while distracted, was not beneficial as Dijksterhuis and colleagues claim, but harmed memory of the attribute dimensions. Nevertheless, intuitive approach in online situations might be beneficial for high-dimensional data. When the decisions were conscious, in the original shopping study, the dimensionality was controlled by asking how many features the subjects were taking into account. Second, attempts to create conscious linguistic descriptions of a face impair intuitive recognition (Schooler & Engstler-Schooler 1990). This implies that the cognitive processes for consciously providing verbal attributes are capacity limited in the way that visual recognition abilities are not. Third, linguistic descriptions of value judgments make the judgments more inconsistent (Levine et al. 1996) and move the decision further from the ones of experts (Wilson & Schooler 1991). To sum up, many ordinary tasks can be challenging for the conscious deliberation and they are better grasped intuitively.

The difficulties in conscious deliberation are believed to result from capacity limits (e.g. Kahneman 2002 based on Nobel price lecture) of working memory. Visualizations are in well-suited position to circumvent these limitations by providing factual input to the process of forming intuitions. High dimensionality entails many independent factors that influence the decisions and high resolution working memory cannot hold much over four chunks for conscious comparisons (Cowan 2000; Zhang & Luck 2008[2]). The capacity limit explanation is supported by the finding that all the college students remembered consciously as much information about the course, but the ones who analyzed the reasons remembered information about less informative details (Wilson & Schooler 1991). Especially, the decisions seem only to be harmed if the person is unaware of the underlying reasons for the intuitive decisions. The above-described findings suggest that intuitive mental processes are not useful for giving justifications or explanations but they can be useful for grasping initial overviews of multi-dimensional data. After exploratory data analysis, the intuitions can be confirmed by conscious reasoning (Tukey, 1977).

The following tentative findings substantiate the suitability of visual interaction for providing input to intuitive overviews of multi-dimensional data. First, it is fair to assume that the most visual processes are automatic and unconscious (see Publ. IV). Therefore, vision is not limited to the subset of conscious cognitive processing. The next section (3.2) provides more direct evidence of this. Second, visual interaction through gaze shifts takes place in time-scales of a third of a second and the cognitive processing related to that time-scale are mostly non-conscious (Ballard et al. 1995; 1997; Publ. IV). This is also the time-scale that is essential for visual searching of details from data-intensive visualizations (Ch. 5 & 6). Third, users of interactive visualizations are often not conscious of all the logical foundations of their visually grounded assumptions. For instance, the users of the election service (Ch. 2 and Publ. I and II) were mostly not knowledgeable about the underlying algorithmic simplifications. They were simply interacting with visual relations. Similarly, the users of other graphical interfaces (e.g. word processing) were shown to utilize rapid on-line inferences, instead of consciously remembering common labels or learning the effects of typical functionalities (Payne et al. 2001). Fourth, visual interaction was shown to be important for learning and problem solving in studies with electronic circuits (Cheng 1999) and real pulley systems (Ferguson & Hegarty 1995) when compared to static figures. Interactive visualizations can function similar to these real systems. These reasons make vision and visual interaction suitable for grounding intuition on multi-dimensional data.

How the interaction has been described in this section as a basis for intuitions is in many ways similar to the original motivations to develop graphical user interfaces (GUI) (Sutherland, 1963; Schneiderman, 1981; 1997;

---

[2] An infamous historical suggestion for capacity limit was “the magical number seven” but even the inventor, Georg Miller suggested that his rhetorical findings about the number seven were likely to be a coincidence (Miller 1956).

Hutchins et al., 1985; Frohlich, 1993). They promoted: (i) continuous visual representation of the object of interest, (ii) rapid, reversible, and incremental actions, and (iii) direct manipulation of objects, instead of command language (Shneiderman, 1981; 1997). However, the challenge for implementing those three objectives was the missing cognitive and behavioural analysis (Shneiderman, 1981). This thesis faces that challenge.

In the field of visualization, the interaction as an active exploration was considered essential already when the term *information visualization* was invented (Robertson et al. 1989). In the initial model, movement of a hand provided input for the computational system, which in turn provided visualization as output. From the user's perspective, interaction takes the form of eye-hand coordination (Smith et al. 2000; Johansson et al. 2001; Hayhoe et al. 2003; Land, 2004). Therefore, the thesis emphasizes the active role of users in perception of visualized information. Here, the gaze shifts are treated analogous to manipulations (e.g., with keyboard) because they influence the information that user receives through retinal image. Gibson (1962) argues that both vision and touch, as senses, should be considered as channels for exploration instead of channels for passive receiving. For example, the shape of cookie cutters was better identified by active touch than by passive touch. Similarly, humans understand visual environment often by observing consequences of their manipulations.

Mental representations of complex static visualizations are also constructed progressively, so understanding their construction process can be instructive for the design of interaction as well. The quality of the mental representation of static image depends on the following perceptual processes (Publ. IV):

(i) peripheral vision (Sect. 5.1),
(ii) instantaneous information retrieval using eye movements (Sect. 3.3 & 5.4),
(iii) the use of general knowledge for object recognition (Sect. 5.1),
(iv) rapid conceptual memory (Sect. 5.1),
(v) mental imagery based on longer interval memory.

When interaction is added to static visualization, computers become low-effort drawing and editing aids for constructing diagrams (Rogers 1999). Analogically, drawing sketches often facilitate creative work. The only difference is that in the visualization the content to be understood and conceptualized is not extracted from the mind. The power of computers is in the ease of pre-processing and drawing the data.

This section has linked the interaction to constructing initial explorative intuitions about the data. In addition, such interaction can confirm or disconfirm the initial intuitions. For instance, user might think that his conclusion about a best-matching item is independent of certain dimension and be surprised that the identity of the best match changes, when that dimension is eliminated. The truth-value and generality of an assumption about the data can be evaluated from changes of a useful visualization, when its parameters are interactively manipulated. For instance, similar items in a similarity map should remain close to another. When the initial intuitions are based on visualized data, they become less arbitrary. Visual modality is suitable for this purpose because visual perception typically involves unconscious intuitive processing and multi-dimensional decisions seem to be done better with intuition.

3.2 Vision for executing actions and mental representations

Overview of Section

The second section (3.2) discusses the fundamental objectives of visualization. Typically the objective is defined as supplementing information for mental representation of user's. The thesis suggests a new objective based on analysis of human visual system. Visual input can be used directly to guide actions. Actions include gaze shifts toward relevant content, manipulation of parameters that control the visualization, and other choices mediated by user-interface. The section refers to empirical evidence from visual neuroscience indicating differentiated large-scale neural systems for vision-based mental representation (ventral stream) and vision-based guidance of actions (dorsal stream). The thesis hypothesizes that this general phenomenon of human vision extends also to the use of interactive visualizations. Thus, the users' capabilities to selectively attend to relevant locations, manipulate the way the data is presented, and to observe the consequences of manipulations are important for understanding. The critical factor for non-expert use of such visualizations is often the ability to operate it (e.g., looking at informative locations), as opposed to the amount of immediately conveyed factual information (i.e., absorbing visual details). This realization has important implications for selecting criteria to evaluate compatibility of different visualization types with human visual cognitive processes.

The previous section reviewed empirical findings showing the importance of interaction for building intuitions about the data. Efficient interaction has been characterized in GUI literature (Hutchins et al. 1985; see also Frohlich 1993) by (i) immediate translation of thoughts to physical control actions (execution) and (ii) system output consistent with user's goals (evaluation). This section argues that such interactivity has fundamental implications for what is the role of visualization for human perceiver and what are the criteria for efficiency in visualization. It is often assumed that the role of visualization is to deliver as much data as possible for human vision to be mentally represented (e.g., Kalawasky, 2009). However, this section presents evidence for view that facilitating mental representations is only one of the two goals of visualization. The other goal would be facilitating unconscious vision-based actions and manipulations. The user should know what is in the data, but also how to obtain the wanted pieces of information when needed.

The evidence for the broader *two-objective view* of visualizations comes from analysis of human vision and the underlying neural processing systems. The general consensus of vision researchers is that these objectives are also general objectives of human vision system. This thesis provides support for hypothesis that interactive visualizations would be most efficient if they are adapted to both of these ecological objectives of human vision. The ecological objectives of human visual system are the following. First, visual perception attempts to represent the outside world. Second, human actions are facilitated by the visual system. The modern view of visual neuroscience associates these objectives to different neural systems. This thesis suggests that visualization should not neglect the importance of the latter objective and the corresponding neural system.

The first objective, the representational task, has been taken to amalgamate in David Marr's (1982) famous attempts to explain vision as an inverse computation. *The inverse problem* for the human vision is to separate 3D sources (surrounding geometry, its reflectance, general illumination, and observer's viewpoint) from two 2D retinal image. However, the underlying assumption that 3D representation is constructed has been criticized for not being necessary or consistent with human behavior (Bülthoff et al. 1995). Another problems for the representational account, is that the quality of the retinal image seems not to be good enough for very accurate representation and visual processes are often constructions (O'Regan 1992; Publ. IV).

The second objective, facilitation of actions, views humans as active agents in their environment. Consistently, the next section analyzes cognitive information processing aspects of user's environment and not only those of the user. According to this view, the purpose of vision is to facilitate actions (Kirklik et al. 1993a; b; Zhang 1997; O'Regan & Noë 2001; Noë & Thompson 2002). Instead of passively representing the environment, visual system can acquire information on-demand-basis (Rao & Ballard 1995). Active vision has been characterized as an exploratory activity mediated by *sensory-motor contingencies* (O'Regan & Noë 2001; Noë & Thompson 2002). For instance, motion parallax is a strong depth cue as the depth related contingencies influence the changes in visual image when head moves.

The representation-action dichotomy has reliable support from cognitive neuroscience in the form of distinct processing pathways (ventral and dorsal) that support the mentioned two tasks of vision. This differentiation of visual pathways was first observed in hamsters (Schneider, 1969) and has mostly been studied with monkeys (since Ungerleider & Mishkin 1982). Importantly, these processing streams are also differentiated in humans (e.g., Culham and Kanwisher 2001), and they have different cognitive roles and relate to functionally different brain locations (temporal and parietal lobes). Initially, the differentiation was identified between object recognition and spatial representations (Ungerleider & Mishkin 1982). In the context of information visualization, these spatial representations could be used for content that is initially non-spatial (Gattis and Holyoak 1996). Later, however, the association with spatial aspects of the environment was interpreted to signify actions (Goodale & Milner 1992; Milner & Goodale 1995), especially non-conscious explorations of the environment (Neisser 1994; Norman 2002). This supports the conclusions of the previous section about usefulness of interaction for data-driven intuitions. Orienting to the environment by eye movements or by pointing is related to this processing stream (dorsal and retinotectal pathway), whereas conscious object recognition is related to the other stream. To summarize, the neuroscience review supports the existence of two tasks for vision (representing and acting) but their practical importance might vary with context (Norman 2002).

The existence of a large-scale visual neural system dedicated to actions has important implications for visualization. For the visualization to be understandable and usable, it should also facilitate those actions that reveal the essential structures of the data. By contrast, the effectiveness of visualizations has been evaluated with respect to Marr's model (Kosslyn 1985). Accordingly, the discriminability of details has been suggested as the ranking criteria for different graph types (Kosslyn 1985; Cleveland & McGill 1984; 1985; Zacks et al. 1998; critically in Wainer & Thissen 1988). Yet, some of the same researchers have admitted that "the primary purpose of visualizations is not to convey numbers with as many decimal places as possible" (Cleveland & McGill 1984, p. 535). In contrast, this thesis argues that other aspects than accuracy with details should be used to evaluate the effectiveness. The ecologically more crucial issues are the commonly occurring visual-cognitive subtasks and especially those subtasks that are difficult. The differences in difficult and frequent subtasks are likely to result in larger changes in quality of user performance with visualizations. An example of a particular subtask is reading a label and tracking an identity from a group of lines, bars, or points. The later discussed eye movement recordings constitute an important method to evaluate these subtasks. The most efficient graph type is the one that makes the difficult visual cognitive operations easier.

As a conclusion from this section, both constructing mental representations and facilitating actions are the most fundamental objectives of human vision. It is not yet certain whether these two objectives are also fundamental for the use of visualizations. This is a large empirical question mostly raised in this section due to increased interaction in computer visualizations. Even so, static visualizations can also include actions to be facilitated such as gaze shifts between names of train stations on a city map. At minimum, disregarding the cognitive processing resources related to vision-based actions seems wasteful. The neural differentiation suggests that additional action-related cognitive processes would interfere less with the other representation-related processes than additional representation-related processing would. Therefore, action related processing could be considered as a potential resource for benefiting from visualizations.

3.3 Cognitive tasks in visualization instead of mind

Overview of Section

Based on the view presented in previous section, the third section (3.3) discusses the different ways to minimize the amount of consciously memorized facts. This is often necessary for the understanding due to the mismatch between large data quantities of data and working memory capacities of an inexperienced user. This is achieved through *externalizing* user's mental operations. That is, visualization application as an external cognitive artefact can relieve user from memorization and other information processing. The decisions based on complex visualizations are cognitively more demanding than those based on simpler ones, and therefore the benefits of externalization become more relevant with multi-dimensional data.

One of such benefits is an opportunity to relieve users from some of the otherwise necessary cognitive burdens of mentally representing visualized data. While the previous sections discussed the interaction with external visualizations, this section will focus on the benefit of having the external medium in itself. The benefit of having the medium is externalization of mental processing. Later parts of this section offer many techniques by which visualizations can be designed to externalize cognitive processing load. The section also analyzes the way many of these techniques were used in the political application (Ch. 2).

The general challenge to understanding of multi-dimensional data results from sequential nature of human perception. The ability to receive and processes information instantaneously is limited and the memory-based integration of sequentially received information is difficult. This thesis argues that visualization can reduce the difficulties by externalizing the information processing. Especially, Chapter 4 suggest a way to present structured information from which gaze shifts can be used to access information without much effort. Some eye movement studies (Ballard et al. 1995; 1997) have shown that complex tasks are often serialized and the users index the information with eye movements to minimize the memory load. The observed subjects rarely memorized information from more than one visual item within a fixation or two. This section argues that this spatial structure might replace the need to construct mental representation in human memory. The details need not to be mentally represented if they can be quickly retrieved. Visual interface and user has been argued to form *an actively coupled system* that process information as a whole and both components can influence each other (Clark & Chalmers 1998). For instance, tiles in a game of Scrabble would not merely be visual objects but parts of thinking process as well. This is supported by brain research showing that tools are in many ways mentally represented as own body (Maravita & Iriki 2004). As a consequence the concept-creating human (see Ch. 4) is placed within a larger and more powerful system (than the human alone) that includes also visually and manually operated “resources”.

It is obvious that externalization of long-term memory (e.g., in calendars) is useful, but the following example will show that users of time intensive interfaces, such as interactive visualizations, naturally minimize their memory loads as well. Kirsh and Maglio (1994) have proposed a category of actions with user interfaces termed *epistemic,* as opposed to pragmatic. The studied expert players of computer game, Tetris, frequently made (otherwise) counterproductive actions to simplify the decision task. This was shown to be efficient with strict time limits because the mental rotation of a piece was slower than the actual rotation of the same piece. Similar simplifying actions have been observed in many real-life problem solving situations (Kirsh 1995).

Cognitive ergonomy forms another aspect of conceptualizing visualizations as simplifications. That is, the perceived content is made more adapted to the cognitive processes of the user. Humans are able to interpret information in very different forms, but they are more experienced with some of the forms. For instance if the task is multiplication, Hindu-Arabic numerals are more suitable than Roman numerals. Compare LXVIII x X to 68 x 10. The perceiver needs to convert or re-represent Roman numerals in Hindu-Arabic forms because modern people are so familiar with 10-based representations (Zhang & Norman 1995; Scaife & Rogers 1996). However, the situation is quite different if the objective is to keep on track how many cars passed by.

The importance of context for human cognition has led some researchers to extend the conceptions from external memory representations (e.g., Hertel 1993; Cary & Carlson 2001) to external cognition or external cognitive artifacts (Larkin & Simon 1987; Hutchins 1995a; b; Scaife & Rogers 1996; Zhang 1997; Clark & Chalmers 1998; Cox 1999; Wilson 2002). Thus, human abilities (e.g., memory capacity) in context of visualizations should take into account the possibilities of environment for storing and processing information. Distributed information processing is better understood with framework of socio-technical system, for instance, when analyzing aviation (Hutchins 1995a) or navigation with large ships (Hutchins 1995b). When human cognition is dynamically coordinated with such socio-technical context, it is called *situated cognition* (e.g. Greeno & Moore 1993; Beer 2000; Clark 2008). This thesis points out that visualizations provide an opportunity for experts and novices from different backgrounds to communicate by referring to parts of visualizations as opposed to difficult concepts that might not be shared.

The observations related to external cognition can be used to extract visualization principles (Larkin & Simon 1987). In the following, some cursory guidelines are extracted from the literature of external cognition to reduce the cognitive loading with visualizations.

(i) Tools are often designed for a particular information-processing task. For instance, nautical maps only represent certain features consistently from three-dimensional globe and compromises many others. The selected projection can either represents constant heading with a straight line (Mercator's projection) or enables estimations of distances on meridians (Robinson's projection).

(ii) Salience and explicitness should correspond to the importance of details (Scaife & Rogers 1996; see also Bauer & Johnson-Laird 1993). Only some part of the information is explicitly represented and they can become explicit only when user looks at the navigation tool in a certain situation. For instance, the political application (Ch. 2, Publ. II) only provided detailed information from one selected location at a time. Visual components can also control memory cues, for instance, temporally hiding some functionalities from a pull-down menu can reduce the complexity of decisions by influencing items in the working memory (Kirsch 1995; cf. affordances in Gibson 1979; Norman 1999).

(iii) Placing the information where it is needed (Larkin & Simon 1987) minimizes the requirements of mental working memory. **(iv)** Some information structures are poorly understood as explicit symbols (Hutchins 1995b). According to Hutchins (1995b p. 118), computation is not only manipulations of symbols, but also "propagations of representational states across the medium", for instance manual operations with tools. Offering these manual operations (e.g., with sliders) can make dynamics of a complicated system more apparent when user observes the consequences (Rogers 1999). **(v)** The existence of visual representation can make possible mental operations that would otherwise be impossible (Chambers & Reisberg 1985). **(vi)** Part of the information processing can be done when the visualization, or a tool in general, is designed. For instance, reading a value from a logarithm table or slide ruler can relieve user from mental arithmetics. **(vii)** The visual form of table makes the number of data points more apparent and indicates the potential missing data by visual gap (Schwartz 1971).

From the perspective of external cognition, the interactively used political service (Ch. 2, Publ. I and II) offers a radically new kind of setting for users to conceptualize multi-dimensional political issues than the traditional questionnaires with concluding lists of best- matches (candidates to vote for). First, SOM algorithm simplifies multi-dimensional relations into two-spatial dimensions. Alternatively, user can attempt to construct the simplified representation mentally, as described in Chapter 4. However, it is likely that users choose arbitrary subset of dimensions that meet their working memory limitations (e.g. four according to Cowan, 2000) and the understanding of the given information is limited. Second, these memory limitations are not as severe for conceptualization in the novel approach because the spatial structure is externalized from user's memory to interactive contingencies (cf. Rao & Ballard 1995; Ballard et al. 1997). These contingencies were already discussed in the previous section with the action-based explanation of human vision. In short, some visualized structures of the data are expected to remain constant irrespective of manipulations of the control parameters of the visualization. Third, the process of manipulating these parameters and observing consequences itself can externalize cognitive operations. For instance, user might question whether the similarity between two candidates is based on the opinions about a certain issue and simply manipulate the choices to test this hypothesis. The comparable mental operations are likely to be very difficult.

3.4 Summary of contributions

The modern needs for information visualization were argued (in Berg & Kojo, 2011) to result (a) from large data quantities (a historically fundamental challenge that has only escalated) and (b) from multi-dimensionality (as wish to integrate or combine different data types). The general solution of the past two decades has been the use of interactive visualization tools and this section has pointed three ways to make these tools understandable. First, it is argued that visualization should focus on supporting exploration and unconscious accumulation intuitions about the data instead of supporting conscious and rational inferences. This idea stems partly from the review in Publication IV about the delusions related to visual perception, and the conclusion that visual perception is in fact mostly not conscious.

Second, the importance of observing consequences from manipulative actions was recognized during the visually guided exploration process. Section 3.2 (also in Publication IV & VII) reviewed the findings of visual neuroscience concluding that there is a notable neural system specialized for visual guidance of motor actions and observation of the following consequences. The review implies a novel conclusion that interactive visualizations should not only strive to facilitate user's accurate mental representation of the data, but also facilitate learning of visually guided manual skills to retrieve the data.

The third way to make complex representation understandable is to relieve cognitive capacity by externalizing storing and processing of information from users' minds. The mentioned actions (redirecting gaze) for rapid retrieval of information provide important example of this. This idea has been recognized only recently for visualization (Liu et al. 2008), so more extensive review of the externalization possibilities was provided in Section 3.3 and in Publication V. Again, it was explained how the application of Chapter 2 incorporates many of these possibilities. Altogether, this chapter answered the first research question (of Sect. 1.2) by arguing that understanding of complex visualizations benefit from unconscious learning, interactivity, and externalization.

4 CONCEPTUAL SPACES AS STARTING POINT FOR PREPROCESSING

Overview of Chapter

This chapter answers the third research question (of Sect. 1.2):

What kinds of mental representations are thought to be central to human concept

formation and the commonly used dimension-reduction algorithms?

The purpose is to provide a starting point for the challenges such as those encountered with the development of the application described in Chapter 2. In what form should one represent information about abstract concepts such as 'EU-supporter' or 'EU-critic'? The represented candidates could have different degrees of 'EU-supporting'. The main statement of Gärdenfor's (2000) Conceptual Spaces (CS) theory is that such abstract concepts are mentally represented in very much the same way as the fundamental low-level perceptual dimensions, such as colors.

CS theory generalizes several properties of low-level categorization to abstract categorization. *Categorization,* in this context, is simply a consequence of being forced to associate something to one of the mutually excluding sets. For instance, the voters can only select one candidate to vote for. CS theory argues that the conceptual categories are defined by the following:

(i) distance metrics of continuous spatial representation with few dimensions,

(ii) dimensions are *reduced* from multi-dimensional information about the world, and

(iii) category membership can be fuzzy and depend on similarity to internal prototype.

These characteristics of concepts are explained in more detail in separate Sections (4.1 - 4.3), but they are first introduced by contrasting them with a possible (but not only) alternative. First and most importantly, conceptual categories are spatial as opposed to being logical (e.g., propositions by Pylyshyn, 1986). For instance, low-level percept of auditory pitch is represented so that the dimension orders pitches in space of cortical substrate (Romani et al. 1982; Pantev et al. 1989; Howard et al. 1996). The logical ordering would not require such spatial mapping. Second, if the number of relevant dimensions is too high the spatial mapping is done only after reducing the dimensionality. For instance, the number of odor receptor types (~ 1000) far exceeds the feasibility of direct spatial mapping (Buck & Axel 1991). A mentioned alternative for data-driven dimension reduction would be to choose only a subset of subjectively appealing dimension, for instance choosing car according to favorite color. Third, category membership is a matter of degree, even though external forces sometimes guide humans to dichotomous decisions. Not every color called 'blue' is perceived as equally blue. We will come back to this issue in Section 4.3. Later discussed fuzzy sets and prototype theory are means to address this issue.

The three characteristics of CS theory are based on vast number of empirical studies related to low-level perception. Some of them are based on earlier theories (e.g., prototype theory in Sect. 4.3). The following three sections will review more empirical studies to support the decision to emphasize these characteristics in visualization. Gärdenfors (2000) objectives were more focused on linguistics than vision. In addition, Publication III provides an example of process model for these theories and that way supports the feasibility of the theories. The important new contribution of the CS theory is to extend these findings from perception to more abstract conceptual categories (e.g., in politics) and the evidence for that is much more limited. The available evidence is discussed next.

4.1 Continuous space in human topographies

The most concrete evidence of continuous mental spaces comes from neural topographies. In topographical organization, the position in neural surface indicates the value on ecologically important perceptual dimension. At least the following dimensions have topographical organization: (i) auditory frequencies or pitches (*tonotopy*: Romani et al. 1982; Pantev et al. 1989; Howard et al. 1996), (ii) stimulated locations on skin (*somatotopy:* Nakamura et al. 1998; Maldjian et al. 1999; Blake et al. 2002), and (iii) several visual features (e.g., location of projection to retina, *retinotopy:* Engel et al. 1997; Hadjikhani et al. 1998). The above-mentioned topographies represent simple perceptual features, but there are also topographic representations for visual object categories (Haxby et al. 2000; 2001). Perhaps, these experiments provide the first evidence for spatial representation of abstract conceptual categories as suggested by CS theory. In these cases the dimensions have general ecological importance and the number of perceptual dimensions do not exceed that of physical space (3 or less).

The dimensions of these topographies are called *quality dimensions* in Gärdenfors' CS theory. The distance in terms of quality dimensions of CS is used to answer the question about the similarity of perceived instances. Quality dimensions are subjective internal structures that organize representations of perceptual input content. However, not every quality dimension can be directly extracted from sensory receptors. Some need to be constructed using several sensory dimensions. For instance, there is no one receptor for 'richness' of red wine.

4.2 Human dimension reduction

Overview of Section

When the constructed dimensions are discussed the following research question (of Sect. 1.2) is evaluated:

Can human concept formation processes alone simplify multi-dimensional perceptual

input from external world analogously to dimension-reduction (DR) algorithms?

The underlying practical objective for this question is to evaluate the human potential for understanding visual representations based on constructs of statistical algorithms. If humans can do the analogous job compared to DR algorithms, they have better chances of being able to understand something from the output of these algorithms. Moreover, the output similar to human conceptual structures is perhaps easier to understand.

Peter Gärdenfors' (2000) Conceptual Spaces (CS) theory argues that it is, indeed, possible to find analogies between human concept formation and discussed dimension reduction algorithms (Sect. 2.3). The argument is supported by empirical examples of this section. The examples deal with natural data sets of (i) olfaction, (ii) visual object categories, (iii) tactile sensation, and (iv) taste. This thesis characterizes these findings as dimension-reduction because the (input) sensory representation has higher dimensionality than that of conceptual quality dimensions. This is taken as support for the Conceptual Spaces theory.

The first example comes from a study of olfaction (Buck & Axel 1991) based on which Nobel prize was awarded in 2004. Low-dimension space of odor sensations is constructed by projecting 1.000 different types of randomly placed receptors to topologically organized olfactory bulb (Axel 1995). The similarities of (up to 10.000 different) scents are subjectively represented with much fewer dimensions than 1.000; hence the dimensions are reduced.

The second example describes construction of new representational dimensions. This happens when one learns to recognize that the previously unfamiliar face belongs to an acquaintance. The neural activity measured by fMRI (Edelman et al. 1998) or by single cell recordings (Kiani et al. 2007) represents visual objects in space similar to MDS output.

The third example from tactile sensations reveals subjective quality dimensions constructed from several different kinds of receptors under the skin. Using MDS, the data was collapsed from pairwise similarity ratings (of surfaces under subjects' index fingers) into three dimensions (Hollins et al. 1993). Adding more dimensions would have not accounted much more variance. The three dimensions were associated with verbal labels: (i) roughness-smoothness, (ii) hardness-softness, and (iii) springiness. Nevertheless, many of the verbal descriptions did not correlate much with the quality dimensions. This also supports the ideas of intuitive processing of multidimensional data that were discussed in Section 3.1. A possible fourth example could come from suggested taste spaces (Churchland 2002, p. 293; based on Bartoshuk & Beauchamp 1994) but evidence for that is less conclusive. The above examples show that humans do reduce dimensionality analogous to DR algorithms.

4.3. Protypes and fuzzy sets for separating categories

Overview of Section

The previous sections (4.1 & 4.2) have described the conceptual spatial representations, but have not dealt with the metrics of similarity or how these spaces can be used for categorization. The person interested in 'rich' tasting red wine might want to categorize tastes belonging to this category as opposed to other red or white wines. In CS, categories are associated with regions surrounded by prototypes, but prototypes have been discussed much before CS theory.

A *prototype* is a representation of typical features of its members and feature here refers to particular values on the dimensions. Prototypes are often confused with some typical member of that category (as in Gärdenfors, 2000, p. 84) but, in fact, prototype is an abstract representation of "central tendencies of categories" (Rosch & Mervis, 1975, p. 574). The mistake is understandable because the idea of prototypes was originally derived from observation of actual typical members. The first empirical studies of this sort of typicality involved color categories (Berlin & Kay 1969; Kay & McDaniel 1978; Kay & Kempton 1984; Kay et al. 1991; but see Saunders, 2000). The categories were found to have the following properties. First, categories were shaped by their use instead of language. Second, centers of gravities (prototypes) were identified for distributions of each of the colors. Third, members that were closer to the centers of gravities were learned faster, later categorized faster, and used more in novel situations.

These core findings of the prototype theory were also interpreted in terms of fuzzy sets (Kay & McDaniel 1978). *Fuzzy set theory* (Zadeh, 1965) provides empirically motivated constraints for prototype representation (Zadeh, 1982), which can be used to construct an actual process model for implementing the categorization, as in Publication III. Prototype theory itself does not provide any kind of a process model (Rosch & Mervis, 1975; Rosch 1978). Fuzzy set theory assigns a degree of prototypicality for each data entity (e.g., tasted wine) if it lies within a given bound from the prototype. This way fuzzy set theory can provide constraints for distance metrics of later-developed CS. However, many practical questions are left open for actual implementations. One of such questions is deciding how many prototypes and concepts are needed. For instance, Russian language differentiates light and dark blues leading to more sensitive color discrimination compared to English speakers (Winaver et al., 2007). When CS theory is used for decision-making, the question of categorization is converted into question of which prototype is most similar to the data point to be categorized.

The notion of choosing among prototypes has several consequences. First, all members of categories are not equally prototypical or equally "good representatives" of the category. Applying fuzzy set theory to CS leads to graded similarity measures as opposed to binary categorization. Second, categorization is more difficult at the borders between the categories. As a consequence, perceptual system has tuned to be more sensitive at the borders (Harnad, 2003), for instance between color categories blue and green (Kay & Kempton, 1984). Third, the categories are developed as alternatives for one another. In CS theory, this notion takes form of *contrast classes*. For instance the skin colors beige and brown are linguistically referred to as 'white' and 'black'. Gärdenfors' (2000, p.121) idea is that 'white' and 'black' are opposite ends of continuum in conceptual space of skins. Fourth, Gärdenfors argues that conceptual system strives toward such spatial representations, in which categories are associated with convex regions. Convexity can simply be a consequence of interpolating between prototypes. Convexity implies "betweenness", which is to say that if two points belong to a category so do points between them. The possible convexity of categories is not easy to evaluate empirically and it is not as essential for current purposes as prototypicality. Overall, prototype theory leads to mental representation, which relates perceptions to one another instead memorizing them separately (as in exemplar theory of Medin & Schaffer, 1978).

4.4 Summary of contributions

This chapter discussed the fundamental properties of CS theory in relation to supportive empirical evidence. One reason for the new empirical evidence is that they are related to the visual perception, where as Gärdenfors was more interested in promoting the view for linguistic purposes. The motivation of visual perception was that information visualizations preprocessed to conform to the fundamental properties of CS are assumed to be easier to comprehend. For instance, if conceptual similarities of political candidates were mentally determined in continuous space, then spatial visualization, which emphasizes the continuum, would be easier to understand. The general means to develop the political application (Ch. 2) towards spatial representation was to utilize SOM dimension reduction and prototypes. One special way to emphasize the graded nature of conceptual similarity was interpolating gradient color shading between the opposite prototypes.

The new contribution of this chapter was to relate the CS theory to empirical findings and to outline its fundamental properties to be emphasized in visualization. Moreover, Publication III provided an example of a process model based on CS theory to support its feasibility. The empirical evidence supported affirmative conclusion for the second research question as humans were shown to be able to reduce dimensionality of multi-dimensional sensory input. They were simply not choosing few dimensions from input, but instead constructed new ones. The third research question about the nature of mental representations was answered by outlining the fundamental properties of CS theory. Those were spatial representations, dimension reduction, and categories separated after interpolating between prototypes.

5 VISION RESEARCH AS STARTING POINT FOR DRAWING VISUALIZATIONS

Overview of Chapter

The previous chapter discussed the form in which preprocessed multi-dimensional data would be understandable. This chapter discusses properties of human visual system to determine how that information can be delivered through visual modality efficiently. We move from the question of what to represent to the question of how to do the representing. An important challenge for the representing is caused by the narrow extent of high-acuity vision (see Fig. 1), which makes it difficult to locate the relevant items from data-intensive visualizations. This challenge corresponds to the fourth research question (of Sect. 1.2):

What kind of visual representation of the preprocessed data would guide user's eyes and attention to the most relevant locations?

Before describing any vision research or visualization solutions, the related objectives for understanding multi-dimensional data need to be discussed. A good starting point for the research of visualizations is to focus on the aspects of their use that are most difficult. Large quantities of multidimensional data often lead to difficulties in locating and integrating details. Integration, as relating spatially separate details, has been argued to be the most essential, difficult, and time-consuming task with visualization (Carpenter & Shah 1998; Gattis & Holyoak 1996; Guthrie 1988; Guthrie et al. 1993). Consistently, Publication V suggests that the integration task is suitable task for differentiating usable visualizations from difficult ones by using behavioral experiments. These difficulties have been argued to result mainly from spatial separation of details (Sweller et al. 1990), which is discussed next as the increased separation leads to the use of peripheral vision. This section suggests ways to minimize these difficulties based on vision research.

A similar objective is to guide user's attention instead of capturing it. Guidance takes place contingent on user's objectives. This leads to different cognitive processes in control of attention and eye movements (Posner 1980; Müller & Rabbit 1989; Berger et al. 2005; Shipp 2004) compared to capture. There are two main reasons to avoid attention capture, for instance by movement (Boyce & Pollatsek 1992; Hillstrom & Yantis 1994; McCrickard et al. 2001) or appearing object (Todd & Van Gelder 1979; Yantis & Jonides 1984; 1990 Faraday & Sutcliffe 1997). The first reason is that designers of the visualizations typically do not know the exact motivations of the users. The user might want to ignore the emphasized piece of information and focus on others. In that case, the capture is distracting (Gillie & Broadbent 1989; Czerwinski et al. 2000; Oulasvirta & Saariluoma 2004). Second, different dimensions of multi-dimensional data should be treated equally by representing visual features (Publ. V), except for possible data-driven preprocessing (Sect. 2.3). For instance, using famous Chernoff's (1973) faces instead of sector diagrams to represent dimensions of political questionnaire data (Publ. II) would not conform to this objective, because the mapping of dimensions is arbitrary. Visual system is innately paying more attention to the dimension visually represented by eye shape (Emery, 2000) at the expense of dimension represented by nose shape.

Thus, guidance is not simply an issue of increasing saliency, but perhaps equalizing the saliency across dimensions (van den Berg et al., 2008) and facilitating search and selective focus between prominent aspects of the data (Publ. VI). Spatial attention and central vision need to be guided, because they are often needed for effective perceptual performance. Without them even large changes in visual field might not be detected, as in change blindness and inattentional blindness experiments (Neiser 1979; Rensink et al., 1997; Mack & Rock 1998; Rensink 2000; Simons & Rensink 2005; Simons & Silverman 2004; Henderson & Castelhano 2005).

In general, the vision research and suggestions for visualization in this chapter are selected based on the proposed sequential framework of interacting with visualizations (Fig. 1). The first section (5.1) approaches the challenges of finding visual features for data representation by reviewing the use of peripheral vision in behavioral studies. Peripheral vision is shown to be important for several visual tasks. Its importance seems to be in providing rapid visual information in more abstract level than what is obtained later by central vision and focused attention. The behavioral studies also give indications about the proper spatial scale for peripheral representations.

The rest of the vision research sections involve visual searching, because the visualization interface is not easy to use if one has to spend much time consciously searching for the relevant details. The sections are divided by the conditions and definitions of search targets to (i) single feature targets, (ii) multi-feature targets, and (iii) elaborate searches with several eye movements. *Targets* are simply the items that fulfill the perceptual search criteria. For instance dots or squares are targets when looking for a certain city on a map. The two objectives of the reviews are to identify visual design features (e.g., color coding, spacing between items) that determine the performance level and to outline the methodology for studying graphs based on how human visual processing has been studied.

5.1 Peripheral vision in guidance of attention

A suggested general solution to aid the search and integration is to provide outlines of data through peripheral vision. According to the sequential framework of interacting with visualizations (Fig. 1), the main purpose of outlines is the guidance of attention instead of immediate delivering detailed information to the user. The outline helps the locating task by restrict spatial attention and eye movements to the most likely regions, where the searched detail is located. Visible outlines can only result from large elements due to the resolution in peripheral vision. The outlines presented in peripheral vision provide context for the detailed information the user is inspecting. This is reminiscent of fisheye views (Furnas, 1981; 1986), in which important locations or aspects are visually magnified surrounded by general context. Human vision acts analogously when the gaze location selects an important location to be the target of eye fixations. The important location is "magnified" as the visual acuity increases after the eye movement. Later the outline helps the integration task by providing a superstructure for working memory to relate the details. Publication II utilized the above ideas, but also special solutions for the search and integration difficulties. The object displays (Ch. 6) were used to represent all the dimensions close to one another and the similar items were placed closer than dissimilar items with help of the SOM algorithm. The peripheral outlines in that SOM visualization were created by sketching fuzzy boarders between categorical regions (Fig. 3) as in CS theory (Sect. 4.3).

The benefits from the above solution are supported by the empirical findings discussed next. First, the peripheral vision is shown to help in very many situations including searching. Second, the findings show that peripheral vision is in position to guide central vision, because it seems to operate faster than central vision. This is crucial for time-intensive interactions with visualizations (Ch. 3). Accordingly, multi-dimensional visualizations should be evaluated based on perceptibility of overviews (Publ. V). In an alternative model, the perception of details precedes and the overview is constructed from the details (Marr, 1982). If that would be true, the psychophysical evaluation of visualizations should focus on discrimination of fine details (Kosslyn 1985; Cleveland & McGill 1984; 1985; Zacks et al. 1998; critically in Wainer & Thissen 1988). Third, the findings provide indications for what is the proper size for peripherally presented visual elements.

Several lines of behavioral research have shown that peripherally seen global patterns are used in directing central vision (Publ. IV). First, memory studies with rapidly presented images have shown that general large-scale concepts (e.g., picnic) can be extracted from image before eyes have time to move (Potter 1975; 1976; 1993; Potter et al., 2002). The overview seems to be remembered, but not the details (Irwin 1996; O'Regan 1992; Intraub 1997). Second, recognition of detailed objects have benefited from large backgrounds, again, before eyes have time to move (Biederman 1972; 1982; Biederman et al., 1973; 1974; 1992; Boyce et al. 1989; Boyce & Pollatsek 1992). These two findings generalize to natural images but also to line drawings, typically used in visualizations. Third, neurophysiological findings have shown that visual processing of larger parts precedes processing of smaller details (Bullier & Novak 1995). Fourth, groups of items have been compared faster than their parts (Pomeranz 1981; 1989; Ariely 2001; Chong & Treisman 2003). This is probably due to perceptual grouping into pre-attentively perceived textures (Julesz 1981; Bergen & Julesz 1983), as used in visualizations (Healey et al. 1996; Healey 1998; Healey and Enns 1999; Interrante 2000). Finally, visual searches became slower when eye movements are monitored and the peripheral vision concurrently occluded or filtered (Rayner & Fisher 1987; Watson et al. 1997; Bertera & Rayner 2000; Loschky & McConkie 2002; Reingold et al. 2003 Geisler et al. 2006, see also modeling by Najemnik & Geisler 2005; Torralba et al., 2006 and tunnel vision in Luo et al. 2008). Altogether, these findings suggest that peripheral visual information is often used to guide centrally perceived details.

The usefulness of outlines in peripheral vision depends on the perceptibility and the impact of the outline as guiding factor relative to currently perceived information at central processing. For instance, seeing faces or reading labels automatically slow down disengagement of attention and gaze, even if they are not capturing stimuli. Publication VI and the filtering studies cited above give indications for the minimum size at given distance, so that the outline is perceived in first place. The size that has most impact can be estimated from experiments in which objects (e.g., letters) of different size provide conflicting information (Navon 1977; Kinchla & Wolfe 1979; Antes & Mann 1984). Those experiments corroborate the general conclusion of this section that large objects up to the size of 6° – 9° (degrees visual angle) influence the perception of smaller ones, more than vice versa. Consistently, HCI researchers have avoided using visual elements larger than 5° (Casey & Wickens 1986; Tullis 1997).

5.2 Visual search for deviant single feature

Overview of Section

The objectives of the following three sections are to (i) develop methodology for using the visual search tasks with information representing graphs, and (ii) identifying relevant visual design features that affect the efficiency of the visual search. Therefore, this section provides the first and general part of the answer to the last research question (of Sect. 1.2):

How should multi-dimensional visualizations be evaluated quantitatively
by psychophysics?

The discussion has two objectives: (i) to provide motivation and support for the experimental methods and stimuli, used in Publication V and VI, and (ii) to determine the design features that make visual searching efficient.

The reviewed experiments were designed to reveal properties of visual processing. The applied experiments in Publication V and VI, in contrast, were designed to evaluate stimuli capable of representing multi-dimensional information. Thus, there was no direct motivation to replicate, contradict, or to discover new phenomena as in the reviewed experiments. Rather, the applied research aim to extend the generality of basic research findings to settings that model the real use of visualization applications.

Another objective of this chapter is to determine the design features that make visual searching efficient. The conditions for the search are typically divided according to common criteria into three sections (5.2 - 4). The first of these sections (5.2) focuses on such situations in which a single deviant feature (color, size, shape etc.) can be used to detect a deviant data value. This has been called *feature search*. Publication VI provides an example of how this knowledge can be used to design a visualization from which outliers can be located very quickly. The following sections provide a more general discussion about the used low-level visual features, because the psychophysically evaluated graphs only utilized orientation, length and size. However, sometimes several features need to be evaluated together in order to detect a deviant or searched individual from multidimensional data. Searches related to simultaneous perception of multiple features are discussed after that in Section 5.3. These have been called *conjunction searches*. These two sections describe the easiest cases of locating interesting individuals and data dimensions. In other cases, users need to perform several eye movements (Sect. 5.4) and extensive analysis of details using central vision.

The efficiency is typically associated with *parallel* or *preattentive* visual processing. Search is called parallel if the performance level does not deteriorate from adding spatially distinct objects or simultaneously processed features (in conjunction search). The visual design features, which have been shown to influence single feature locating abilities, are listed next with short definitions (Geisler & Chou 1995; Carrasco et al. 1995; 1998; Palmer et al. 2000; Bertera 2000; Levi 2008).

Key results of Section

Single feature search depends on:

(i) ***Target/non-target similarity*** – difference between target and non-target in a visual feature,

(ii) ***Decision criterion*** – threshold for visual feature of an item to fulfill target criterion,

(iii) ***Speed-accuracy trade-off*** – tendency to delay response for more accurate responding,

(iv) ***Eccentricity*** – distance between an item and the center of the gaze, and

(v) ***Crowding*** – adjacent items harm identification if their spatial distance is too small.

Only the crowding and eccentricity are further discussed next because designers of visualizations can adjust them independent of data and users by changing the distance between the visual objects. Crowding harms object recognition if the distance is too small. The required distance increases linearly with the eccentricity (Bouma 1970). Consequently crowding is a problem of peripheral vision. However, crowding should not influence detection of the item's presence, but only its identification based on several integrated visual features (Levi 2008; Pelli et al. 2004 Pelli & Tillman 2008) - the conditions discussed in next section. Publication VI showed first indications of crowding phenomenon with graphs, but more research is needed to evaluate the real significance of crowding for the information visualization. In fact, crowding could be beneficial for the texture visualizations, when object recognition is not the main task. Again, this underlines the importance of developing evaluation criteria for the efficiency of visualizations, which is a major objective of this thesis.

The methodological question of how to measure the efficiency of search from graphs could also be developed based on visual search experiments. The most important factor to be controlled is the eccentricity (cf. visual acuity in Fig. 2) because the eccentricity of an object depends on the gaze direction.

Otherwise the performance level is determined by an irrelevant random factor, namely where the test subject's gaze happens to be when the task begins. There are two general approaches for this problem. The first visual search experiments in 1960s measured the eye movements (Williams 1966). Alternatively, most of the studies 1980s and 1990s have measured manual reaction times after eliminating the eye movements by short presentation times. The latter approach is much more easier and was followed by a seminal study of Treisman & Gelade (1980). The criterion for the search efficiency was the lack of *set-size effect,* meaning that the manual reaction time should remain constant when additional non-target items are included. However, this thesis suggests that eye movement measurement should be used for evaluating visualizations if this is not too inconvenient. The choice of using either eye or hand as response modality might have important consequences for the related visual neural processing, as discussed in Publication VI. Eye movements are ecologically important according to the proposed sequential framework of interacting with visualizations (see Fig. 1). Eliminating eye movements by an artificial time is not representative of real perceptual dynamics and interactions with visualizations (Chmiel 1989; Klein & Farrel 1989). Gaze shifts constitute a better model for the real use compared to manual reaction times because gaze shifts are related to information retrieval by high-acuity central vision and they are a precondition for effective manual operations (Ballard et al., 1995; Land & McLeod, 2000; Land & Hayhoe, 2001; Johansson et al. 2001; Hayhoe et al. 2003; Land 2004), especially with the pointing devices (Smith et al. 2000). The time needed for manual operation with a novel object might even triple if the gaze shifts are prevented. Thus, eye movements recordings are need to reveal the efficiency of graphs in interaction sequences longer than half a second.

For the shorter visual perception tasks, this might not be necessary. When user initiates a change in a dynamic interface the initial gaze point is often known. For instance, in the political service discussed in Chapter 2, when user selects one alternative, the gaze point surely is at the alternatives (see Fig. 3). Then, the eccentricity of the alternatives is small, and the eccentricity of the SOM view is larger. The exact eccentricity depends on the viewing distance and the size of the (window on the) display. The selection updated the SOM view and, for instance, a too fast onset would have been interruptive for inspection of the alternatives. The possible interruption would have been caused by an attention capture by motion or onset at the SOM view located in visual periphery. Attention capture closer to gaze point is likely to be less harmful as the eccentricity, and the corresponding visual acuity, do not change that much. Thus, the designer can often temporarily reason where the gaze is located.

5.3 Visual search for conjunction of several deviant features

As mentioned, many useful possibilities for multi-dimensional visualization emerge if more than one visual feature can be used to represent the different dimensions. More subtle difference can be represented and the relevancy of an item might depend on more than one visual feature. The unwanted consequence is that user's ability to differentiate the relevant items can become worse. For instance, the conjunction search for unique targets might require selecting red and small item, instead of simply choosing small item. The following discussion evaluates the extent to which the search difficulty increases under different conditions with the additional items.

The conjunction search related conclusions from vision research have changed in past few decades and the visualization principles should change accordingly. In 1980s, conjunction search was generally thought to happen serially (Treisman & Gelade 1980; Treisman 1982), which would make data-intensive visualization difficult to use. However, the generality of this conclusion has been denied in many studies after that (e.g., Luria 1975; Nakayama and Silverman 1986; Chmiel 1989; Wolfe 1994; van den Berg 2008). Nevertheless, the textbooks for the design of visualizations have advised not to several features in conjunction (e.g., Bertin 1967/1983) except for special conditions (three exceptions in Ware 2000, p. 170).

The reason could be that the efficiency of conjunction search in complex visualizations is difficult to predict. It depends on the following visual design features:

Key results of Section

Multi-feature search depends on:

(i) adding a redundant feature[1,2]

(ii) density of items[3,4]

(iii) eccentricity[5,6]

(iv) perceptual grouping[7,8]

(v) target/non-target similarity[9,10]

(vi) visual variation in non-target items[10,11]

1 = Wolfe et al. 1989, 2 = Treisman & Sato 1990, 3 = Cohen & Ivry 1991, 4 = Julesz 1981, 5 = Carrasco et al. 1995, 6 = Geisler and Chou 1995, 7 = Pomerantz 1981, 8 =Treisman 1982 9 = Quinlan & Humphreys 1987, 10 = Treisman & Gormican 1988, 11 = Duncan & Humphreys 1989

The last two are related to the obvious result that target saliency improves the performance (Treisman and Gormican 1988) because fewer objects need to be considered (Rayner and Fisher 1987). The above studies are sampled from very large number of conducted studies (see e.g., Wolfe 1998) based on their relevance for visualization.

However, the conjoint effects of the above-identified design features for search efficiency are difficult to evaluate using the existing vision research literature (but see van den Berg et al. 2008). Tentative implications of the early findings can only be given. First, the spatial distance is important for layouts (Kirch 1995) and can be used efficiently with colors. Colors can limit the number of considered alternatives (Egeth et al. 1984) up to twenty items in four colors (Duncan 1989). Second, shape can guide gaze to the most informative location alone (Williams 1966; Viviani & Swensson 1982; Findlay 1997; Publ. VI) or in conjunction with colors (Findlay 1997), even if immediate target identification is not possible. This again supports the earlier conclusion that peripheral vision is utilized more in guidance of central vision and not in recognition of fine details.

5.4 Browsing by moving eyes

This thesis has repeatedly emphasized that even static visualizations often convey the information sequentially to the user due to eye movements. Therefore, the ease of serial eye movements should also be considered with efficient designs. By contrast, eye movements have typically been recorded for conditions characterized by serial search (see Rayner 1998 for review). Eye movement records of otherwise efficient designs can reveal the efficiency of peripheral processing because small temporal differences can accumulate compromising the usability of visualizations. Even an increase on the scale of milliseconds can influence the chosen strategies by diminishing retrieval of accurate information, which increases the use of fallible subjective memorization (Ballard et al. 1995; 1997; Gray & Boehm-Davis 2000; Cary & Carlson 2001), for instance when programming a VCR (Gray & Fu 2004). With a more positive approach, this thesis has outlined the possibilities of external cognitive artifacts that reduce cognitive load from the user's mind (in Sect. 3.3). Gaze direction is an informative measure for benchmarking suitable visualizations as it reveals the visibility of peripheral items (Jacobs 1986) and the requirements for spatial attention (Williams et al. 1997).

Several findings from vision research indicate that the ease of gaze shifts can lead to performance differences with otherwise efficient designs. First, graphs often include visual indexes (numbers or letters), for which perception depends more on central vision and gaze location than peripheral vision

(Bertera & Rayner 2000). Second, eye movements are useful for retrieving information farther from the periphery (over 8° in Scialfa & Joffe 1998) with larger graphs. Thus, the search can be parallel only within a limited region, and eye movements are relevant for more distant items. The third reason to measure eye movement is that unnaturally limited stimulus presentation time might not predict performance in context of use (Chmiel 1989; Klein & Farrel 1989). The worst case would be that eye movement and manual response performance would indicate different types of visual features and graphs to be optimal for perceptual system. Eye movements have been associated with intuitive and unconscious processing (sect. 3.1), where as manual responses with conscious object detection (Norman, 2002).

Based on previous sections, peripheral vision can either speed up reaction times (i) by providing information about the target or (ii) by guiding gaze toward it. The eye movement recordings support the latter alternative as gaze can selectively directed based on the visual features used in reaction time experiments (Luria & Strauss 1975; Findlay 1997; Williams & Reingold 2001; primates: Motter & Belky 1998a; Bichot & Schall 1999a; 1999b). The selectivity refers to the probability of looking first at the task-relevant target item (or one of them) as opposed to any other item. The gaze is shifted to increase visual acuity and performance is not deteriorated by small inaccuracy of eye saccades (Motter & Holsapple 2007; Motter & Belky 1998a, exp. 4; Viviani & Swensson 1982; Findlay 1997). Very accurate saccades might not be useful as they require more time (according to Fitts' law; Fitts 1954; Ware and Mikaelian 1987; MacKenzie 1992). If the visual acuity at the target location is not sufficient after the saccade, an additional corrective saccade is performed.

The accuracy of visually guided saccades and the need for corrective saccades depend on the following visual design features:

Key results of Section

Search with gaze shifts depends on:

(i) distracting items close to the target location[1]

(ii) density of those items[2]

(iii) eccentricity[3]

(iv) target/non-target similarity[4,5]

(v) visual variation in other items[6]

(vi) hurry to move eyes[6,7]

1 = Walker et al. 1997, 2 = Motter & Belky 1986b, 3 = Findlay 1997, 4 = Jacobs 1986, 5 = Gilchrist at al. 1999, 6 = Zelinsky 1996, 7 = Hooge & Erkelens, 1999

As a rule, the factors that determine the accuracy of eye movements are similar to the factors that determine the speed of manual responses (see previous section).

5.5 Summary of contributions

In short, this chapter dealing with vision research lead to three different kinds of conclusions:

(I) Peripheral vision helps more in guidance of gaze than in directly conveying data.

(II) Visual factors summarized in above three result-boxes determine locating abilities.

(III) Experiments with visualizations should control those factors, timing, and gaze point.

This chapter approached the task of selecting visual features to represent the data from the perspective of proposed sequential framework (Fig. 1) and the challenges of finding and integrating details from data-intensive visualizations. The proposed method to reduce these challenges is using peripheral vision to convey outlines. Even though the use of preattentive features has been discussed in visualization literature (Bertin 1967/1983, Ware 2000), the vision research related to peripheral vision and visual searching have, to my knowledge, not been discussed thoroughly before. More importantly, the identified importance of outlines challenges the traditional psychophysical evaluation criterion that emphasizes the discrimination of details. In spirit of the sequential framework, outlines were not so much intended to represent as much data as possible, but to lead users toward the most useful details. The discussed empirical findings gave reliable support for the usefulness of peripheral vision and for its use as a guiding factor due to the speed of peripheral visual processes. Further indications were given about the suitable size of visual features in periphery. In short, the visual features should be larger and coarser (low spatial frequencies) than those representing the detailed information.

The rest of the chapter evaluated search task of vision research for choosing the proper visual features to represent data. The vision research results do not warrant conclusions about the implications for changing one visual feature in real and complex visualization applications. Instead, this chapter focuses on identifying the visual design features that are critical for the ability to locate details. The critical features were discussed separately for the situations that do not require gaze shifts and those that do. The cases in which target details are define either by one feature or several features were also differentiated. In general, the search performance depends on the similarity between target and non-targets, variation in non-target items, eccentricity, density of items, and perceptual groupings. In addition, when several features need to be used concurrently the performance depends on the redundant features and possible visual crowding. Vision research provides exact limits for eccentricity at which crowding starts to harm object recognition. The spacing between the items should be larger when they are positioned further from the central vision. The ability to shift gaze accurately toward the target depends very much on the same features than the manual reaction times. The importance of gaze shifts for visual acuity and hand-eye coordination were emphasized.

Aside from providing guidelines for the development of visualizations, vision research findings were utilized for developing evaluation methodology for graphs. An ecological valid criterion for data-intensive visualizations is the ability to locate the relevant detail. This thesis suggests adapting vision research methodology to compare human performance with graphs - components of visualization application – in controlled laboratory settings. The next chapter and Publication VI report examples of search experiments for visualization purposes. In such experiments, the identified visual features, which determine the search performance, should be controlled or their contribution evaluated. Eccentricity, distance from gaze point to search target, was identified as the most important determinant for that. Eye movement measurement and restricted presentation times were discussed as alternatives.

A typically disregarded implication of locating ability is the aid for constructing integrated memory representation from perceived details. Visual memory processes are not needed as much if the time between looking at several items is shorter. This is argued to be important for understanding complex visualizations.

6 NEW STUDIES OF LOCATING AND INTEGRATING ABILITIES WITH GRAPHS

Overview of Chapter

The primary purpose of Chapter 6 is to prove that the proposed criteria for effectiveness of multi-dimensional visualizations can be operationalized as quantitative experiments. This chapter answers the last research question (of Sect. 1.2):

How should multi-dimensional visualizations be evaluated quantitatively

in psychophysical experiments?

The reported experiments involved two tasks: search and integration.

The psychophysical studies reported in Publications mostly involve two kinds of experiments, different versions of searching and integration. In the search experiments, subjects' task was to locate an item with the deviant feature or identify that deviant feature from a set of alternatives. In the integration experiments, subject evaluated similarity of different graphs based on shapes that depicted several data values on different dimensions. It is called integration as all the values were encoded into short-term visual memory at once.

The integration task combined two principles from the sequential framework of understanding visualizations. The first principle involved sequentiality as details (20 values on different dimensions) were combined, and then compared based on visual memory lasting over an eye saccade. In contrast, earlier experiments have focused on accuracy of perceiving individual values instantly. Another principle relates to the idea that comparisons are performed intuitively on outline shapes. The subjects utilized the individual values conjointly in a way that they were not able to verbalize with any rules. The locating experiment involved three principles of framework. First, the task response required task-motivated filtering of information by selective gaze shifts toward items and dimensions. In contrast, many visual search tasks have required a simple manual response for detecting presence of a visual feature.

Second, peripheral vision had to be used for guidance of those gaze shifts. Third, the presentation time for the peripheral information was shorter than time needed to initiate gaze shifts. Thus, successful performance would not indicate any delay in browsing of visualization due to peripheral information processing. This was argued to be important criterion for externalizing multi-dimensional details into spatial layout. In opposite case, a mentally straining strategy of memorizing details could be favored to speed up the browsing. Then, memorization would reduce mental resources available for evaluating the perceived information.

Next, the main aspects of the methods are only shortly summarized, but more details can be found from Publications VI and VII. The objectives of the following experiments differed from the vision research of previous chapter as the experiments evaluated representations of data instead of evaluating human perceptual processes. There were three general objectives for the experiments. First, knowing the perceptual aspects of using sector and bar diagrams uncover the contribution of sector diagrams to the popularity and inferred usability of the discussed political visualization service. The two sector diagrams in that service were the only user interface components, which were not evaluated qualitatively in the prototype development phase (Publ. I). Second, speed and ease of perception in the experiments reveal effectiveness for visualization use in more general. The special objective of the experiments was to evaluate these diagrams for clarifying the dimension reduction visualizations. To ascertain the generalizations from the particular political task domain, the experiments eliminated the entire political context except the numeric data structures. Third, meaningful results would indicate that the developed evaluation criteria, locating and integrating, are useful (operationalization). Thereby many other diagrams could be evaluated accordingly. When the above objectives were not conflicted, the methods of the experiments were mimicking basic research as much as possible in order to support generalizations from the results.

6.1 Integration study

Aims and Methods

Objectives of the integration study led to methodological differences compared to basic vision research. First, and most importantly, graphs were used as stimuli instead of the simplest visual features optimal for receptive characteristics of human brain (e.g., typical Gabor patches). Graphs are representations because the visual features (shape, size, color etc.) are intended to refer to some other aspect of reality (possibly even hypothetical aspect) than the visual feature itself. So far, such well-controlled studies of visualizations have been scarce (see Publ. V for a review). The utilized graphs, sector and bar graphs, were chosen among a larger set of alternatives because they are easier to compare than the dimension reduction algorithms, and they do not pose (much) artificial linking of different data points. The set of alternative graphs consisted of object displays, which utilize shape for relating and integrating several data values (Casey & Wickens 1986; Carswell & Wickens 1987; Wickens & Hollands 1999, p. 125-129; Bennet et al. 2000; Bennet & Walters 2001). This objective made them suitable for multi-dimensional data because integration facilitates processing of several data dimensions simultaneously. The benefit of object displays is that different dimensions are represented by visual features of a single object instead of the corresponding features in several objects. The underlying relations between the dimensions and the mapping from data to visual shapes have depended on the choices of the researchers. For instance, subjects might try to multiply some values of dimensions with triangle shape (Carswell & Wickens 1987). The following experiments utilized much more dimensions (25) than the earlier experiments (two or three).

Second, the represented information was randomly sampled from the election data (opinions of candidates). The simple artificial data used in basic vision research might not be suitable for visualization research because it might lead to overemphasizing the distinctiveness of cluster structures, which would not be representative of real use. Nevertheless, Publication VI also tested simplified data as perceptual stimuli using sector diagrams.

In the procedure of the first experiment, subjects evaluated the holistic similarity between graphs representing 25 data points each (Fig. 5). The task was to select the graph that was perceived as more similar to the standard from two alternatives. The novel evaluation criterion, integration, was used because it influences the ability to understand visualizations more than discrimination of fine details (Carpenter & Shah 1998; Gattis & Holyoak 1996; Guthrie 1988; Guthrie et al. 1993). Most of the earlier experiments have measured thresholds for discriminating small differences (critically in Wainer & Thissen 1988). Most notably Cleveland and McGill (1984; 1985; Cleveland 1984) have based their ranking of graphs on fine discrimination. The similarity evaluation was based on visual short-term memory of a graph perceived only during 100 ms. This models the real perceptual sequence by allowing the subjects to see the stimulus for as long as they typically would during a single browsing fixation. The time is shorter than required for typical saccade (200 – 300 ms), controlling the eccentricity effects (Ch. 5) as a fixation marker controlled the gaze position. Performing accurately under this constraint entails no delay for normal gaze sequence when visualization is browsed. The objective of Experiment 2 was to ensure that the new integration criterion and the above condition would not result in too much compromise with the traditional criterion.

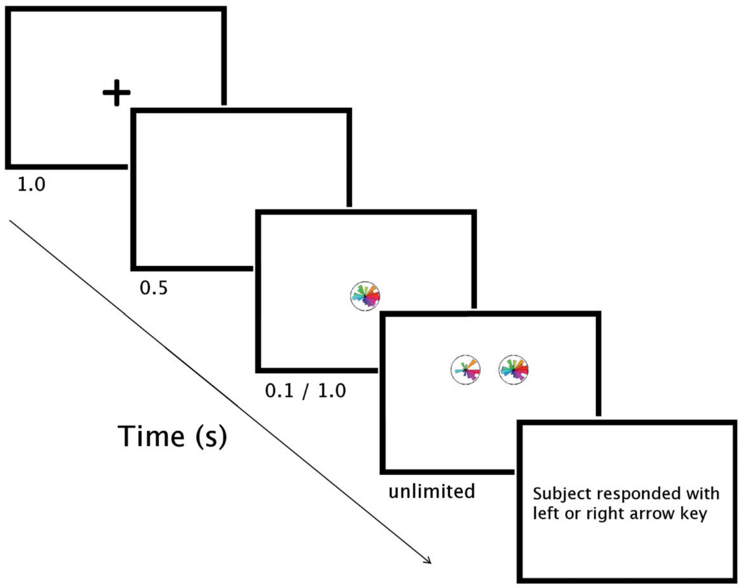


Fig. 5 Procedure of Experiment 1 in Publication V. The phases are illustrated for sector diagram in order of presentation (from top left to bottom right): fixation cross, blank screen, standard diagram, alternative diagrams, and subject's response.

Thus, the experiment was controlled to measure reaction time (RT) and accuracy of choices as dependent variables in two-alternative-forced-choice (2AFC) setting. The manipulated independent variables of the Experiment 1 were similarity of the alternatives, type of graph (sectors or bars), color (B/W or colors), and presentation time (100 ms or 1.0 s). In addition, the interactive effects of these factors on dependent variables were analyzed. In Experiment 2, the task was to detect differences in details. No colors were used and the presentation time was not limited.

Results and discussion

In the integration experiments (Exp. 1 in Publ. V), the performance accuracy was very high with both graph types and depended mostly on the size of the differences in data values. In addition, the performance with sector diagrams was consistently better than with bar diagrams. Sector diagrams also resulted in lower discrimination thresholds for individual data values than bar diagrams (Exp. 2 in Publ. V). Thus, the trade-off (between discrimination and integration) typical for object displays (Wickens & Hollands 1999, pp. 125-129) was not observed in these experiments.

The following reasons for the advantages of sector diagrams were suggested. First, locating a diagram from a group of similar diagrams is easier with sector diagram because they can be perceptual segmented using conspicuous gap between them. The gap results from round overall shape. Second, each dimension is easier to identify because the identity of a sectors is also redundantly represented by orientation and not only by location. Orientation is a fundamental visual feature, for which brain has specially tuned cells (Hubel & Wiesel 1959), and its use could distribute the neural representation for larger group of cells. Nevertheless, the findings are also encouraging for the use of object displays in general. Users can remember and compare information rapidly across gaze shifts when they are browsing multi-dimensional data points represented using different object displays.

6.2 Search study

Aims and Methods

The objective of Publication V was to model perception of sector graphs during a single fixation. The complementary role of Publication VI is to model the perception and the gaze sequence between the fixations. However, as mentioned not much information is received during the saccades, and therefore, the real objective of Publication is to evaluate how well users can locate critical information from large sets of graphs.

Earlier parts of this summary (and Berg & Kojo, 2011) have argued that the main challenge with multi-dimensional visualizations results from the large number of data entities and dimensions. Thus, the search, as a model of real use, requires locating the relevant entities and dimensions. In object displays, a single entity or an individual corresponds to a single graph. The objectives of Publications VI and VII correspond to Ullman's (1984; Rao & Ballard 1995; see also Ballard et al. 1998) two primary visual routines for obtaining abstract properties from visual field. First, the identification tasks (Publ. V and Exp. 2 & 4 in Publ. VI) correspond to the routine of extracting properties from a pointed location. In line with the experiments, Ullman especially discusses the importance of bounding contours for the comparing shapes of different component parts. Second, the localization tasks (Exp. 1 and 3 in Publ. VI) correspond to the routine of attending or pointing to specific aspects of environment. In this thesis, the latter routine was examined in basic research set-up of visual search. Together, these routines have been respectively related to the two visual neural processing streams for 'what' and 'where' (or 'how') information (Rao & Ballard 1995), as discussed in Section 3.2.

The reasons for the different search conditions (Exp. 1-4) with sector diagrams in Publication VI are explained next. **Experiment 1** measured the time needed to locate sector diagram (representing a data entity) having a deviant value, and **Experiment 2** measured the time needed to identify that sector (representing a dimension). **Experiments 3** and **Experiment 4** measured eye movements, instead of manual reactions, as the response modality replicating Experiments 1 and 2 respectively. Both the choice of task (identification / localization) and the choice of response modality (manual / gaze) have been argued (Milner & Goodale 1992; 1995) to be differentially related to the visual processing streams (ventral / dorsal). Therefore, all the combinations of these alternatives were evaluated.

In isolation the studied tasks are categorized as feature searches (reviewed in Sect. 5.2), but combining the tasks in Experiment 4 resemble conjunction searching based on two features (Sect. 5.3). Whether the particular task involves conjunction search is relevant because that indicates the risks of crowding effects, as mentioned. Perhaps, the search component involved would be rejected as conjunctions according to strict standards, but the subjects do have to identify the diagram using one visual feature (length) and sector by using another (orientation). The connection between basic vision research and meaningful use of visualizations is often not straightforward on the most detailed level of explanation.

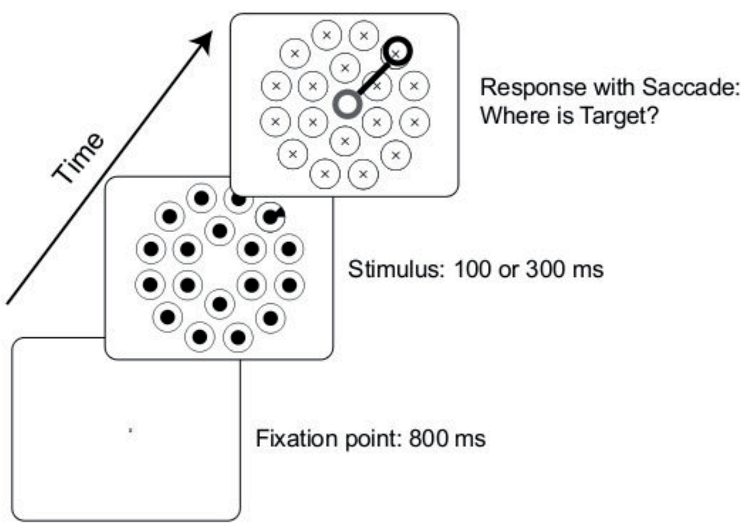


Fig. 6 Procedure of Experiment 3 of Publication VI. Experiment 3 used same diagrams as Experiment 1, but the experiment included sixty diagrams (only 18 presented here for clarity) and measured saccadic accuracy.

As the number of data entities and the size of the entire visualization increase, the eye movements become more relevant for the search. Thus, Experiments 3 and 4 were conducted. This time the experiments measured saccadic accuracy instead of keyboard reaction times (Fig. 6). As a result, the measure becomes more sensitive because there were more diagrams or sectors to look at (480-AFC, Alternative Force Choice) than there were keys to press (2-AFC) in Experiments 1 and 2. Furthermore, manual response times could underestimate the capabilities of users of visualizations, unlike eye movement measures (e.g., Chmiel 1989; Milner & Goodale 1995; see Sect. 3.2). Eye movements are also ecologically more valid response modality than manual keyboard responses because they lead to higher visual acuity, aid the use of pointing devices (Smith et al. 2000), and potentially externalize the visual memory load (Ballard et al. 1995; 1997).

Results and discussion

The visual search experiments were designed to evaluate the identification performance from object displays when the number of diagrams and dimensions increase. The experimental task of the search experiments assumed that the difference between the deviant search target and the non-target items was large enough to be differentiated from peripheral vision. This assumption was confirmed by Experiments 1 and 2 of Publication VI. The main purpose of the search study was to find out if this perceived difference leads to differentiation of the corresponding item and dimension by gaze shifts. This was also confirmed by the Experiments 3 and 4 of Publication VI. The identification performance was independent of the total number of diagrams (up to 60), which is important for data-intensive visualizations. The subjects could identify the deviant diagram from 60 alternatives each having 8 possible dimensions. Consequently, the subjects were able to gaze at correct source from the 480 alternatives presented only for 100 ms. The ability to identify the source proves that the preconditions were filled for externalizing the memory to rapid saccades that access the information when needed.

The localization and identification performance were very accurate in the conducted experiments. Interestingly, performance was worse on the two outer rings than on the two inner rings of experiments in exact accordance with Bouma's (1970) criterion of peripheral crowding. Bouma’s crowding limit is defined according to spacing between items and their eccentricity. However, these experiments were not directly designed to evaluate the crowding and thus it was confounded with lowered acuity at peripheral locations. Thus, the agreement with Bouma's criterion could also be a coincidence. This can be tested with further experiments by compensating the lowered accuracy by scaling the objects larger at the periphery (according to M-factor, Virsu & Rovamo 1979; Anstis 1998; Duncan & Boynton 2003). In a nutshell, the reviewed visual search literature (Ch. 5) outlined some important visual design features (such as eccentricity, target-distractor similarity, and crowding) that should be analyzed when evaluating the efficiency of different graphs. The results are likely to generalize to other visualizations when the saliency of the mentioned features is high relative to other features (listed in Ch. 5). Most importantly, target items should not be too similar with the other items or otherwise they cannot be differentiated.

The contributions of this thesis to information visualization are certainly not limited to the conclusions that can be directly drawn from the results of conducted experiments. More can be gained from the analysis of how to design visualizations, what criteria to use for evaluation, and how to control the evaluation. These analyses were based on a framework of interactive visualization, in which understanding is a sequential process that depends on the input from peripheral vision and the conscious capacities relieved by externalization and use of data-driven intuition. The framework was based on reviewed multi-disciplinary findings from neuroscience, psychophysics, other behavioral research, and related modeling.

Together the results from all the conducted experiments lead to the conclusion that the studied object displays (bar and sector graphs) fulfilled the suggested criterion well. First, these object displays allowed simultaneous attending to all dimensions of one data entity (e.g., product or person). This result suggests that those graphs are suitable for sequential similarity comparisons of very many discrete data entities. This way, object display can complement dimension-reduction and convert some of the important validity, usefulness, and reliability assessments to assessments of visual similarity. For instance, a user doubting the dimension reduction can simply evaluate whether similar data points are placed close to one another. However, those graphs might not be efficient for all the possible multi-dimensional tasks, for instance evaluating the relations between the dimensions per se might be difficult. Second, the experiments showed that the data entities could be browsed within minimal gaze fixation duration (< 100 ms). For sector diagrams, this finding was also shown to extend to searching and identification of deviant data entities and dimensions. The natural gaze sequence with these diagrams is therefore not delayed and the visualization can facilitate externalization of details. The externalization here means that users need not to memorize the details, but instead, they can rapidly access them by spatial gaze point indexing. To summarize all the experiments, the psychophysical experiments (Publ. VI & VII) support the ease of search and integration of information from object displays. Thus, these object displays could effectively help users to overcome some of the challenges in the understanding representations of dimension-reduction algorithms.

7 General discussion

7.1 Proposed principles

Overview of Section

The following principles (P1, P2, & P3) are suggested based on more detailed discussion in Chapters 3 and 4. The first two principles introduce general possibilities to help users of visualizations to understand more complex relations from multi-dimensional data than before. This summarizes how the first research question (of Sect. 1.2) was answered in Chapter 3. The principle three proposes more specific methods for the same purpose and integrates the answers to the next three research questions addressed in Chapters 4 and 5. The last two research questions are not related to any design principles because they involved evaluation of already constructed visualizations. The related answers are summarized in Conclusions based on Chapters 5 and 6.

P1. Visual representation can be designed to relieve user from otherwise necessary cognitive loading.

This principle has been referred to as *external cognition*. The idea is that visualizations not only help users, but they also perform some of the necessary cognitive functions. Thus, information processing is distributed to a larger system than the human alone (see Fig. 1.). The distributed cognition in multi-user systems was not included into the discussion (e.g., in navigation of a ship; Hutchins, 1995a). The external cognition pioneers are from the field of cognitive science (e.g., Hutchins 1995a; 1995b) but the idea of externalization has recently been recognized also in the field of information visualization (Liu et al. 2008). This thesis developed those ideas further by analyzing the different in which visualizations can relieve cognitive burden of their users. As a result, more cognitive capacity is left for the executive decisions compare to the cases when multi-dimensional information is processed without the help of visualizations, either relying on other perceptual modality (e.g., conversations), long-term memory, or purely symbolic visual representations (e.g., reading alphanumeric tables). Multi-dimensionality and large quantities of data only emphasize the limitations of cognitive capacities, and the limitations are more pronounced for non-expert user-group (e.g., Koedinger & Anderson 1990; Ericsson 2003; Kalakoski 2006).

Two principal ways of externalization were distinguished: (a) external memory and (b) interaction functionalities that externalize processing of visualized data. An important example of former is when eye movements index rapid (~ 300 ms) information retrieval based on the spatial location (Ballard et al. 1995; 1997). In that case, user needs not to hold the information in memory but gaze sequence can act as virtual memory book (Rensink 2000).

As a result, many of the limitations of human working memory capacity are circumvented. This possibility also motivated the conducted visual search experiments (see Sect. 6.2) as the external memory resource would not be utilized if memory retrieval takes too long. An example of externalizing processing is the use of dimension reduction algorithms to supplement concept formation in user's mind. Another example is rotating a piece in game of Tetris instead of mentally rotating it (Kirsch & Maglio 1994).

P2. Visualizations adapted to human vision serve both mental representations and actions.

For the reasons above (P1), visual interaction should utilize (a) unconscious and intuitive processing as opposed to logical processing and (b) the visual processing for guiding actions and observing consequences of those actions (manual or gaze). These objectives need not to compete with the more traditional objective of consciously recognizing visual features (shapes, colors etc.) from visualizations. The first part of this principle argues that simple decisions are better done with conscious logical processes and complex multi-dimensional decision with unconscious processes (Dijksterhuis et al. 2006; Newell et al. 2009). In contrast to limiting the objectives of information visualization for building mental models of the visualized data (Kalawsky 2009), the second part argues that learning interaction skills with user interfaces should not be forgotten. The skills refer here to user's ability to perform proper actions that reveal the wanted information for processes of visual perception. For example, user should know what buttons to press and where to look without the need of conscious reasoning. Furthermore, the existence of separate visual processing stream for interactive manipulations and purposeful gaze shifts provides also evidence for the ecological importance of interaction skills (Sect. 3.2). Some aspects of visualization are simply easier to perceive as consequences of own actions. Accordingly, the externalization principle (P1) states that representational demands can be reduced if information is actively revealed when needed. The importance of this principle has increased as the amount of interactivity in visualizations has increased.

P3. Simplifications of dimension-reduction (DR) algorithms can guide user by providing outlines of data through peripheral vision.

This principle is grounded on the answers to three of the discussed research questions (2–4 in Sect. 1.2). The third principle differs from the first two principles because it includes a specific suggestion for the design of multidimensional visualizations. There is no pre-existing relation between dimension reduction algorithms and peripheral vision. However, their compatibility is suggested. First, humans were shown to be able to reduce dimensionality of multi-dimensional perceptual input analogous to DR algorithms. The analogy was evaluated by aligning the dimensions underlying subjective similarity assessments with MDS constructed dimensions. For instance, subjectively constructed dimensions (e.g., smoothness of surface) for tactile sensations when touching surfaces (e.g., leather and tile) (Hollins et al., 1993). The conclusion was that the results of dimension-reduction algorithms could be understandable to users of multi-dimensional visualizations because they themselves construct analogous mental representations.

Second, Conceptual spaces theory was used to outline the information structures that are natural for human understanding: (i) space of quality dimensions with internal distance metrics, (ii) dimension reduction to construct the quality dimensions, and (iii) prototypes as centers for the alternative categories.

Third, behavioral studies of human vision show that peripheral vision is capable of guiding the later gaze sequence of high-acuity central vision. This is efficient perceptual strategy for humans because large visual shapes presented in peripheral vision are typically perceived before smaller details in central vision. Therefore, this principle (P3) could be called as *peripheral outline preview (POP) principle*.

However, not every visual representation is equally perceivable in peripheral vision. Therefore, *visual search* experiments of perceptual processes were reviewed (Ch. 5) to identify the visual design features (e.g., size, spacing, colors) that determine the capability of peripheral outline in guidance of gaze direction. In addition, the neurophysiological limitations of visual system were shortly discussed. This is necessary because, to my knowledge, implications of peripheral visual system have not been thoroughly discussed for information visualization. The typical approach provides only examples as support for suggested principles (e.g., Bertin 1967/1983). The reviewed basic research studies about human capabilities, however, are not sufficient to determine the efficient characteristics of real visualizations because visualizations typically utilize several visual features together. Thus, the previous chapter consisted of psychophysical studies of perceiving graphs.

7.2 Conclusions

The thesis presents an empirically-motivated framework for understanding interaction with visualizations of multi-dimensional data as a sequence of gaze shifts and manual operations. Three principles were derived from that framework to guide the design of visualization:

(i) dimension reduction can provide outlines for peripheral vision to guide gaze shift,

(ii) externalizing and intuitive processing can relieve capacities for decisions,

(iii) observing consequences of own actions can uncover relations between dimensions.

The first principle had the strongest empirical evidence as support. That allowed the most specified conclusions for the design. The empirical findings provide much certainty about the visual features that determine effective use of peripheral vision. The empirical evidence about human dimension reduction and understanding of algorithmic dimension reduction was very much limited. Therefore much less certainty was given to the conclusions about what to represent in periphery as an outline. Similarly the conclusion about intuition in decision-making warrants more uncertainty. A special form of intuitive interaction was specified; observing consequences from own actions. The converging neuroscience findings provide much support for its importance to humans in general. The practical impact of the last two principles for real information visualization has not been directly evaluated (but see Hutchins, 1995a; b). However, the discussed examples from cognitive psychology provide indications that externalization can improve performance with tasks that are likely to be relevant for understanding of visualizations. The research challenge with externalization is to know how to implement it. The benefits of externalization and intuitive processing depend on the cognitive capacity limits, the usefulness of external tools, and their compatibility with perceptual processes. The cognitive capacity limits are a concern because they can restrict the effectively used dimensionality and was shown to cause suboptimal decisions. This is especially important for the users who are not familiar with dimension reduction methods.

Externalization is the most general answer offered to the challenges of mental capacity limits that are exceeded by the large and complex data of modern information visualizations. The idea is simply that functionalities of visualizations can reduce the amount of information processing required from the user. The release of mental resources, in turn, can improve understanding of the data and its use for decision-making. The thesis suggests a novel externalization method for design of visualizations according to which the users' gaze shifts act as spatial index for detailed information. This procedure can relieve users from the need to remember many details when the speed of their perceptual information retrieval exceeds the speed of initiating gaze shifts. Otherwise, visual browsing is delayed.

The thesis has suggested that the earlier objectives of visualization, as maximizing the amount of data directly conveyed to the user, should be extended to cover also the skills of interaction. Selective gaze shifts to the most relevant spatial locations of visualization are likely to be important for the user to understand the consequences of self-induced changes in dynamic visualizations. Moreover, human visual neural system is likely to treat gaze shifts themselves as actions, which influence the flow of information through the eyes and should be optimized. Thus, peripheral visual information might best serve user by guiding user's gaze toward the most relevant aspects of data instead of conveying much data in itself.

The framework and the above principles were also used to evaluate ways to make visualizations efficient for large and diverse user-groups. For the purpose, empirically derived evaluation criteria were formulated to assess the effectiveness of visualizations. These criteria differ from those of typical psychophysical experiments that have measured accuracy of perceiving a single value. The novel criteria were selected based on the behavioral studies of human vision and practical studies about the use of visualizations for inferences. The critical subtasks in using complex visualizations are those that are difficult or very frequent for users. The identified critical tasks for complex multidimensional visualizations were:

(a) locating relevant details and

(b) integrating the details

(e.g., values on different dimensions) to form lasting memory representations.

The thesis showed that these critical tasks could be turned into controlled psychophysical experiments measuring real graphs. The evaluation methodology for graphs was derived from methods designed to evaluate perceptual processes. In addition to methods, the reviewed research of perceptual processes suggests several visual features that need to be controlled in case of visual search. Such psychophysical control requires much effort but this is unavoidable for the sake of reliable conclusions. The reported experiments show that the suggested framework and the evaluation criteria are also useful for benchmarking graphical elements. The future work can benefit from the suggested criteria, but it remains to be shown how the methods need to be adjusted for evaluating other graph types, and which graph types lead to the most efficient performance.

References

Anderson, J., Byrne, M., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. Psychological Review, 111, 1036-1060.
Anstis, S. (1998). Picturing peripheral acuity. *Perception*, *27*, 817-825.
Anstis, S. M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research*, *14*, 589-592.
Antes, J. & Mann, S. (1984). Global-local precedence in picture processing. *Psychological Research*, *46*, 247-259.
Ariely, D. (2001). Seeing sets: representation by statistical properties. *Psychological Science*, *12*(2), 157-162.
Axel, R. (1995). The molecular logic of smell. *Scientific American*, *273*(4), 154-159.
Bailey, G. (1993). *Iterative methodology and designer training in human-computer interface design.* New York.
Ballard, D., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*, 723-767.
Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, *7*(1), 66-80.
Thomas, J. J. & Cook, K. A. (Eds.). (2004). *Illuminating the path.* Richland, WA: The national visualization and analytics center.
Bartoshuk, L. M. & Beauchamp, G. K. (1994). Chemical senses. *Annual Review of Psychology*, *45*, 419-449.
Bauer, M. & Johnson-Laird, P. (1993). How diagrams can improve reasoning. *Psychological Science*, *4*(6), 372-378.
Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, *4*(3), 91-99.
Bennet, K. B., Payne, M., Calcaterra, J., & Nittoli, B. (2000). An empirical comparison of alternative methodologies for the valuation of configural displays. *Human Factors*, *42*, 287-298.
Bennet, K. B. & Walters, B. (2001). Configural display design techniques considered at multiple levels of evaluation. *Human Factors*, *43*(3), 415-434.
Berg, M. 2004. Emergent political concepts in explorative interface visualization. Master's Thesis. Helsinki University of Technology.
Berg, M., & Kojo, I. (2011). Informaation visualisointi. In A. Oulasvirta (Ed.), *Ihmisen ja tietokoneen vuorovaikutus* (pp. 155-171). Helsinki: Gaudeamus Helsinki University Press.
Bergen, J. R. & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature*, *303*, 696-698.
Berger, A., Avishai, H., & Rafal, R. (2005). Competition Between Endogenous and Exogenous Orienting of Visual Attention. *Journal of Experimental Psychology: General*, *134*(2), 207-221.
Berlin, B. & Kay, P. (1969). *Basic color terms: Their universality and evolution.* Berkeley: University of California Press.
Bertera, J. H. & Rayner, K. (2000). Eye movements and span of the effective stimulus in visual search. *Perception & Psychophysics*, *62*(3), 576-585.
Bertin, J. (1967). *Semiology of graphs* (W.J.Berg, Trans.). Madison, WI: University of Wisconsin Press.
Bichot, N. P. & Schall, J. D. (1999). Effects of similarity and history on neural mechanisms of visual selection. *Nature Neuroscience*, *2*(6), 549-554.
Bichot, N. P. & Schall, J. D. (1999). Saccade target selection in macaque during feature and conjunction visual search. *Visual Neuroscience*, *16*, 81-89.
Biederman, I. (1972). Perceiving real-world scenes. *Science*, *177*, 77-80.
Biederman, I. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*(2), 143-177.
Biederman, I. & Cooper, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 121-133.
Biederman, I., Glass, A., & Stacy, W. (1973). Searching for objects in real-world scenes. *Journal of Experimental Psychology*, *97*(1), 22-27.
Biederman, I., Rabinowitz, J., & Stacy, W. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, *103*(3), 597-600.
Blake, D. T., Byl, N. N., & Merzenich, M. M. (2002). Representation of the hand in cerebral cortex. *Behavioural brain research*, *135*(1-2), 179-184.
Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 117-178.

Boyce, S. & Pollatsek, A. (1992). Identification of Objects in Scenes: The Role of Scene Background in Object Naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 531-543.

Boyce, S., Pollatsek, A., & Rayner, K. (1989). Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 556-566.

Buck, L. & Axel, R. (1991). A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell*, *65*, 175-187.

Bullier, J. & Nowak, L. G. (1995). Parallel versus serial processing: New vista on the distributed organization of visual system. *Current Opinion in Neurobiology*, *5*, 497-503.

Bülthoff, H., Edelman, S., & Tarr, M. (1995). How are three-dimensional objects represented in the brain? *Cerebral Cortex*, *3*, 247-260.

Carpenter, P. A. & Shah, P. (1998). A model of perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, *4*(2), 75-100.

Carrasco, M., Evert, D. I., Chang, I., & Katz, S. M. (1995). The eccentricity effect: target eccentricity affects performance on conjunction searches. *Perception & Psychophysics*, *57*(8), 1241-1261.

Carrasco, M., McLean, T., Katz, S., & Frieder, K. (1998). Feature asymmetries in visual search: Effects of display duration, target eccentricity, orientation and spatial frequency. *Vision Research*, *38*(3), 347-374.

Carswell, C. M. & Wickens, C. D. (1987). Information integration and the object display: An integration of task demands and display superiority. *Ergonomics*, *30*(3), 511-527.

Casey, E. J. & Wickens, C. D. (1986). Visual display representation of multidimensional systems: The effect of information correlation and display integrality. *Technical Report, University Of Illinois*, *CPL-86-2 / MDA903-83-K-0255*.

Chambers, D. & Reisberg, D. (1985). Can mental images be ambiguous? *Journal of Experimental Psychology: Human Perception and Performance*, *11*(3), 317-328.

Chen, C., & Yu, Y. (2000). Empirical studies of information visualization: A meta-analysis. International Journal of Human Computer Studies, 53(5), 851-866.

Cheng, P. C.-H. (1999). Interactive law encoding diagrams for learning and instruction. *Learning and Instruction*, *9*, 309-325.

Chernoff, H. (1973). The use of faces to represent points in k-dimensional space graphically. Journal of the American Statistical Association, 69(342), 361-368.

Chmiel, N. (1989). Response effects in the perception of conjunctions of colour and form. *Psychological Research*, *51*, 117-122.

Chong, S. & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*, 393-404.

Churchland, P. S. (2002). *Brain-wise*. Cambridge, MA: MIT Press.

Clark, A. (2008). *Supersizing the mind*. New York: Oxford University Press.

Clark, A. & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7-19.

Cleveland, W. (1984). Graphical methods for data presentation: full scale breaks, dot charts, and multibased logging. *The American Statistician*, *38*(4), 270-280.

Cleveland, W. & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, *79*(387), 531-554.

Cleveland, W. & McGill, R. (1985). Graphical perception and graphical methods for analyzing scientific data. *Science*, *229*, 828-833.

Cohen, A. & Ivry, R. B. (1991). Density effects in conjunction search: Evidence for a coarse location mechanism of feature integration. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(4), 891-901.

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87-185.

Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction*, *9*, 343-363.

Culham, J. C. & Kanwisher, N. G. (2001). Neuroimaging of cognitive functions in human parietal cortex. *Current Opinion in Neurobiology*, *11*(2), 157-163.

Curcio, C. A., Allen, K. A., Sloan, K. R., Lerea, C. L., Hurley, J. B., Klock, I. B., et al. (1991). Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. *The Journal of Comparative Neurology*, *312*, 610-624.

Czerwinski, M. P., Cutrell, E., & Horvitz, E. (2000). *Instant messaging and interruption: Influence of task type on performance*. Paper presented at the OZCHI 2000.

Davenport, J. & Potter, M. (2004). Scene consistency in object and background perception. *Psychological Science*, *15*(8), 559-564.
Dijksterhuis, A., Bos, M., Nordgren, L., & Baaren, R. (2006). On making the right choice: The deliberation-without-attention effect. *Science*, *311*, 1005-1007.
Duncan, J. (1984). Selective attention and organization of visual information. *Journal of Experimental Psychology: General*, *113*, 501-517.
Duncan, J. (1989). Boundary conditions on parallel processing in human vision. *Perception*, *18*, 457-469.
Duncan, J. & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433-458.
Duncan, R. O. & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron*, *38*, 659-671.
Edelman, S., Grill-Spector, K., Kushnir, T., & Malach, R. (1998). Toward direct visualization of the internal shape representation space by fMRI. *Psychobiology*, *26*(4), 309-321.
Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(1), 32-39.
Emery, N. J. (2000). The eyes have it: neuroethology, function and evolution of social gaze. Neuroscience and Biobehavioral Reviews, 24, 581-604.
Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and spatial precision of functional MRI. *Cerebral Cortex*, *7*, 181-192.
Ericsson, K. A. (2003). The acquisition of expert performance as problem solving. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 31-86). Cambridge, UK: Cambridge University Press.
Faraday, P. & Sutcliffe, A. (1997). *Designing effetive multimedia presentations*. Paper presented at the SIGCHI '97.
Ferguson, E. L. & Hegarty, M. (1995). Learning with real machines or diagrams: Application of knowledge to real-world problems. *Cognition and Instruction*, *13*(1), 129-160.
Findlay, J. (1997). Saccade target selection during visual search. *Vision Research*, *37*(5), 617-631.
Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, *47*(6), 381-391.
Freud, S. (1923). The ego and the id. New York: W. W. Norton & company.
Frohlich, D. (1993). The history and future of direct manipulation. *Behaviour & Information Technology*, *12*(6), 315-329.
Furnas, G. W. (1981). The fisheye view: a new looak at structured files. Proceedings from Bell laboratories technical memorandum # 81-11221-9, Murray Hill, NJ.
Furnas, G. W. (1986). Generalized fisheye views. Proceedings from Human Factors in Computing Systems CHI '86, New York.
Gärdenfors, P. (2000). *Conceptual Spaces: The Geometry of Thought*. Boston, MA: MIT Press.
Garner, W. R. (1970). The stimulus in information processing. *American Psychologist*, *25*, 350-358.
Gattis, M. & Holyoak, K. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 231-239.
Geisler, W. S. & Chou, K.-L. (1995). Separation of low-level and high-level factors in complex tasks: Visual search. *Psychological Review*, *102*(2), 356-378.
Geisler, W. S., Perry, J. S., & Najemnik, J. (2006). Visual search: The role of peripheral information measured using gaze-contingent displays. *Journal of Vision*, *6*(9,1), 858-873.
Gibson, J. (1962). Observations on active touch. *Psychological Review*, *69*(6), 477-491.
Gibson, J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
Gilchrist, I. D., Heywood, C. A., & Findlay, J. M. (1999). Saccade selection in visual search: Evidence for spatial frequency specific between-item interation. *Vision Research*, *39*, 1373-1383.
Gillie, T. & Broadbent, D. (1989). What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological Research*, *50*, 243-250.
Goodale, M. A. & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, *15*(1), 20-25.
Gould, J. & Lewis, C. (1985). Designing for usability: Key principles and what designers think. *communications of the ACM*, *28*(3), 300-311.

Gray, W. D. & Boehm-Davis, D. A. (2000). Milliseconds matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of Experimental Psychology: Applied*, *6*(4), 322-335.

Gray, W. D. & Fu, W.-T. (2004). Soft constraints in interactive behavior: The case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science*, *28*, 359-382.

Greeno, J. G. & Moore, J. L. (1993). Situativity and symbols: Response to Vera and Simon. *Cognitive Science*, *17*, 49-59.

Guthrie, J. T. (1988). Locating information in documents: Examination of a cognitive model. *Reading Research Quarterly*, *23*(2), 178-199.

Guthrie, J. T., Weber, S., & Kimmerly, N. (1993). Searching documents: Cognitive processes and deficits in understanding graphs, tables, and illustrations. *Contemporary Educational Psychology*, *18*, 186-221.

Hadjikhani, N., Liu, A., Dale, A. M., Cavanagh, P., & Tootell, R. B. H. (1998). Retinotopy and color sensitivity in human visual cortical area V8. *Nature Neuroscience*, *1*(3), 235-241.

Harnad, S. (2003). The Symbol Grounding Problem. *Encyclopedia of Cognitive Science*.

Haxby, J., Ishai, A., Chao, L., Ungerleider, L., & Martin, A. (2000). Object-form topology in the ventral temporal lobe. *Trends in Cognitive Sciences*, *4, 1*.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425-2430.

Hayhoe, M., Shrivastava, A., Mruczek, R., & Pelz, J. (2003). Visual memory and motor planning in a natural task. *Journal of Vision,* 49-63.

Healey, C. (1998). *On the Use of Perceptual Cues and Data Mining for Effective Visualization of Scientific Datasets*. Paper presented at the Graphics Interface.

Healey, C., Booth, K., & Enns, J. (1996). High-Speed Visual Estimation Using Preattentive Processing. *ACM Transactions on Computer-Human Interaction*, *3*(2), 107-135.

Healey, C. & Enns, J. (1999). Large datasets at a glance: Combining textures and colors in scientific visualization. *IEEE Transactions on Visualization and Computer Graphics*, *5*(2), 145-167.

Henderson, J. & Castelhano, M. (2005). Eye movements and visual memory of scenes. In G. Underwood (Ed.), *Cognitive Processes in Eye Guidance*. (pp. 213-235). Oxford, UK: Oxford University Press.

Hertel, P. T. (1993). Implications of external memory for investigations of mind. Applied Cognitive Psychology, 7, 665-674.

Hewett, T. & Meadow, C. (1986). *On designing for usability: An application of four key principles*.

Hillstrom, A. P. & Yantis, S. (1994). Visual motion and attentional capture. *Perception & Psychophysics*, *55*(4), 399-411.

Hollins, M., Faldowski, R., Rao, S., & Young, F. (1993). Perceptual dimensions of tactile surface texture: A multidimensional scaling analysis. *Perception & Psychophysics*, *54*(6), 697-705.

Houde, S. & Hill, C. (1997). What do prototypes prototype? In M. Helander, T. K. Landauer, & P. Prabhu (Eds.), *Handbook of Human-Computer Interaction*. Atlanta, GA: Elsevier Science.

Howard, M. A., Volkov, I. O., Abbas, P. J., Damasio, H., Ollendieck, M. C., & Granner, M. A. (1996). A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. *Brain Research*, *724*, 260-264.

Hubel, D. & Wiesel, T. (1959). Receptive fields and functional architechture of single neurons in the cat's striate cortex. *Journal of Physiology*, *148*, 574-591.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Hutchins, E. (1995). How a cockpit rembembers its speed. *Cognitive Science*, *19*, 265-288.

Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1985). Direct manipulation interfaces. *Human-Computer Interaction*, *1*, 311-338.

Interrante, V. (2000). Harnessing Natural Textures for Multivariate Visualization. *IEEE Computer Graphics and Applications*, *20*(6), 6-10.

Intraub, H. (1997). The representation of visual scenes. Trends in Cognitive Sciences, 1(6), 217-222.

Irwin, D. (1996). Integrating information across saccadic eye movements. *Current Directions in Psychological Science*, *5*(3), 94-100.

Jacobs, A. M. (1986). Eye-movement control in visual search: How direct is visual span control? *Perception & Psychophysics, 39*(1), 47-58.
Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye-hand coordination in object manipulation. *The Journal of Neuroscience*, *21*(17), 6917-6932.
Jonson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, *4*, 71-115.
Johnson-Laird, P. N. (1981). Cognition, computers, and mental models. Cognition, 10, 139-143.
Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature*, *290*, 91-97.
Kaipainen, M., Koskenniemi, T., Kerminen, A., Raike, A., & Ellonen, A. (2001). Presenting data as similarity clusters instead of list. Data from local politics as an example. In C. Stephanidis (Ed.), *Universal Access in HCI Towards an Information Society for All*. London: Lawrence Erlbaum.
Kalakoski, V. (2006). *Constructing Skilled Images*. Ph. D. University of Helsinki, Helsinki.
Karvonen, K. (2000). *The Beauty of Simplicity*. Proceedings from Proceedings of the ACM Conference on Universal Usability, Washington DC.
Kay, P., Berlin, B., & Merrifield, W. (1991). Biocultural implications of systems of color naming. *Journal of Linguistic Anthropology*, *1*, 12-25.
Kay, P. & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Antropologist*, *86*, 65-79.
Kay, P. & McDaniel, C. (1978). The linguistic significane of the meanings of the color terms. *Language*, *54*, 610-646.
Kerminen, A., Raike, A., & Kaipainen, M. (2000). *Self-organizing map browser for database retrieval*. Florence, Italy.
Kiani, R., Esteky, H., Mirpour, K., & Tanaka, K. (2007). Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *J Neurophysiol*, *97*(6), 4296-4309.
Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architechture for cognition and preformance with application to human-computer interaction. *Human-Computer Interaction*, *12*, 391-438.
Kinchla, R. & Wolfe, J. (1979). The order of visual processing: "Top-down", "bottom-up", or "middle-out". *Perception & Psychophysics*, *25*(3), 225-231.
Kirklik, A., Miller, R. A., & Jagacainski, R. J. (1993). Supervisory control in a dynamic and uncertain environment: A process model of skilled human-environment interaction. *IEEE Transactions on Systems, Man, and Cybernetics*, *23*(4), 929-951.
Kirklik, A., Plamondon, B. D., Lytton, L., Jagacinski, R. J., & Miller, R. A. (1993). Supervisory control in a dynamic and uncertain environment: Laboratory task and crew performance. *IEEE Transactions on Systems, Man, and Cybernetics*, *23*(4), 1130-1138.
Kirsh, M. (1995). The intelligent use of space. *Artificial Intelligence*, *73*, 31-68.
Kirsh, M. & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*(4), 513-549.
Klein, R. & Farrell, M. (1989). Search performance without eye movements. *Perception & Psychophysics*, *46*(5), 476-482.
Kleinmuntz, B. (1990). Why we still use our heads instead of formulas: Toward an integrative approach. *Psychological Bulletin*, *107*(3), 296-310.
Koedinger, K. R. & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, *14*(4), 511-550.
Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*, 59-69.
Kohonen, T. (2002). Modeling of automatic capture and focusing of visual attention. *PNAS*, *99*(15), 9813-9818.
Kosslyn, S. (1985). Review: Graphics and human information processing: A Review of Five Books. *Journal of the American Statistical Association*, *80*(391), 499-512.
Laine, S. (2003). *Using visualization, variable selection and feature extraction to learn from industrial data*. Helsinki University of Technology, Helsinki.
Land, M. (2004). Eye movements in daily life. In L. M. Chalupa & J. S. Werner (Eds.), *The Visual Neurosciences*. (pp. 1357-1368). Cambridge, MA: MIT Press.
Land, M., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? Vision Research, 41, 3559-3565.
Land, M., & McLeod, P. (2000). Form eye movements to action: How bastmen hit the ball. Nature Neuroscience, 3(12), 1340-1345.

Lappin, J. S. (1967). Attention in the identification of stimuli in complex visual displays. *Journal of Experimental Psychology*, *75*, 321-328.

Larkin, J. & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*(1), 65-99.

Laurence, S., & Margolis, E. (1999). Concepts: Core Readings. Cambridge, MA: MIT Press.

Levi, D. M. (2008). Crowding--an essential bottleneck for object recognition: a mini-review. *Vision Research*, *48*(5), 635-654.

Levine, G., Halberstadt, J., & Goldstone, R. (1996). Reasoning and the Weighting of attributes in Attitude Judgments. *Journal of Personality & Social Psychology*, *70*(2), 230-240.

Linville, P. W. (1982). The complexity-extremity effect and age-based stereotyping. *Journal of Personality and Social Psychology*, *42*(2), 193-211.

Liu, Z., Nersessian, N. J., & Stasko, J. T. (2008). Distributed cognition as a theoretical framework for information visualization. IEEE Transactions on Visualization and Computer Graphics, 14(6), 1173-1180.

Loschky, L. C. & McConkie, G. W. (2002). Investigating spatial vision and dynamic attention selection using a gaze-contingent multiresolution display. *Journal of Experimental Psychology: Applied*, *8*(2), 99-117.

Luo, G., Vargas-Martin, F., & Peli, E. (2008). The role of peripheral vision in saccade planning: learning from people with tunnel vision. *Journal of Vision*, *8*(14), 25.1-25.8.

Luria, S. M. & Strauss, M. S. (1975). Eye movements during search for coded and uncoded targets. *Perception & Psychophysics*, *17*(3), 303-308.

Mack, A. & Rock, I. (1998). *Inattentional blindness*. Cambridge: MIT Press.

MacKenzie, I. S. (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction*, *7*, 91-139.

Maldjian, J. A., Gottschalk, A., Patel, R. S., Detre, J. A., & Alsop, D. C. (1999). The sensory somatotopic map of human hand demonstrated at 4 Tesla. *Neuroimage*, *10*, 55-62.

Maravita, A. & Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences*, *8*(2), 79-86.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W.H.Freeman & Co, N.Y.

Matthews, T., Rattenbury, T., & Carter, S. (2007). Defining, designing, and evaluating peripheral displays: An analysis using activity theory. *Human-Computer Interaction*, *22*, 221-261.

McCrickard, D. S., Catarambone, R., & Stasko, J. T. (2001). *Evaluating animation in the periphery as a mechanism for maintaining awareness*. Paper presented at the INTERACT '01, Tokyo, Japan.

McFarlane, D. C. (1999). Coordinating the interruption of people in human-computer interaction. *Human-Computer Interaction - INTERACT'99*. (pp. 295-303). The Netherlands: IOS Press.

Milner, A. D. & Goodale, M. A. (1995). *The visual brain in action*. New York: Oxford University Press.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63(2), 81-97.

Motter, B. C. & Belky, E. J. (1998). The guidance of eye movements during active visual search. *Vision Research*, *38*, 1805-1815.

Motter, B. C. & Belky, E. J. (1998). The zone of focal attention during active visual search. *Vision Research*, *38*(7), 1007-1022.

Motter, B. C. & Holsapple, J. (2007). Saccades and covert shifts of attention during active visual search: spatial distributions, memory, and items per fixation. *Vision Research*, *47*(10), 1261-1281.

Murphy, G. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Müller, H. & Rabbit, P. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistence to interruption. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(2), 315-330.

Najemnik, J. & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*, 387-391.

Nakamura, A., Yamada, T., Goto, A., Kato, T., Ito, K., Abe, Y., et al. (1998). Somatosensory Homunculus as drawn by MEG. *Neuroimage*, *7*, 377-386.

Nakayama, K. & Silverman. (1986). Serial and parallel processing of visual feature conjunction. *Nature*, *320*, 264-265.

Navon, D. (1977). Forest before trees: The preceedence of global features in visual perception. *Cognitive psychology*, *9*, 353-383.
Navon, D. (1983). How many trees does it take to make a forest? *Perception*, *12*, 239-254.
Neisser, U. (1976). *Cognition and reality*. San Francisco: WH Freeman.
Neisser, U. (1979). The control of information pickup in selective looking. In A. D. Pick (Ed.), *Perception and its development. A tribute to Eleanor J Gibson*. (pp. 201-219). Hillsdale, NJ: Lawrence Erlbaum.
Neisser, U. (1994). Multiple systems: A new approach to cognitive theory. *European Journal of Cognitive Psychology*, *6*(3), 225-241.
Newell, B. R., Wong, K. Y., Cheung, J. C., & Rakow, T. (2009). Think, blink or sleep on it? The impact of modes of thought on complex decision making. *The Quaterly Journal of Experimental Psychology*, *62*(4), 707-732.
Noë, A. & Thompson, E. (2002). *Vision and mind: Selected readings in the philosophy of perception*. Cambridge: MIT Press.
Norman, D. (1999). Affordance, conventions and design. *Interactions*, *6*(3), 38-42.
Norman, J. (2002). Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches. *Behavioral and Brain Sciences*, *25*, 73-144.
North, C. (2006). Toward measuring visualization insight. IEEE Computer Graphics and Applications, 26(3), 6-9.
O'Conaill, B. & Frohlich, D. (1995). *Timespace in the workplace: Dealing with interruptions*. Paper presented at the CHI '95, New York, NY.
O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: the world as an outside memory. *Canadian Journal of Experimental Psychology*, *46*(3), 461-488.
O'Regan, K. & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*, 939-1031.
O'Shea, R. P. (1991). Thumb's rule tested: Visal angle of thumb's width is about 2 deg. *Perception*, *20*(3), 415-418.
Oulasvirta, A. & Saariluoma, P. (2004). Long-term working memory and interrupting message in human-computer interaction. *Behaviour & Information Technology, 23*(1), 53-64.
Palmer, S., Vergehese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*(10-12), 1227-1268.
Pantev, C., Hoke, M., Lutkenhoner, B., & Lehnertz, K. (1989). Tonotopic organization of auditory cortex: Pitch versus frequency representation. *Science*, *246*, 486-488.
Pashler, H. (1994). Dual-task inference in simple tasks: Data and theory. Psychological Bulletin, 116(2), 220-244.
Payne, S. J., Howes, A., & Reader, W. M. (2001). Adaptively distributing cognition: A decision-making perspective on human-computer interaction. *Behaviour & Information Technology, 20*(5), 339-346.
Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. Philosophical Magazine, 2, 559-572.
Peli, E. & Geri, G. A. (2001). Discrimination of wide-field images as a test of a peripheral-vision model. *Journal of the Optical Society of America A - Optics and Image Science*, *18*(2), 294-301.
Peli, E., Yang, J., & Goldstein, R. B. (1991). Image invariance with changes in size: The role of peripheral constrast thresholds. *Journal of the Optical Society of America*, *8*, 1762-1774.
Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: distinguishing feature integration from detection. *Journal of Vision*, *4*(12), 1136-1169.
Pelli, D. G. & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience,* 1129-1135.
Perer, A., & Shneiderman, B. (2009). Integrating statistics and visualization for exploratory power: From long-term case studies to design guidelines. IEEE Computer Graphics and Applications, 29(3), 39-51.
Pomerantz, J. R. (1981). Perceptual organization in information processing. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. (pp. 141-180). Hillsdale, NJ: Lawrence Erlbaum.
Posner, M. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3-25.
Potter, M. (1975). Meaning in visual search. *Science*, *187*, 965-966.
Potter, M. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(5), 509-522.

Potter, M. (1993). Very short-term conceptual memory. *Memory & Cognition*, *21*, 156-161.

Potter, M., Straub, A., Rado, J., & O'Connor, D. (2002). Recognition memory for briefly presented pictures: The time course of rapid forgetting. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1163-1175.

Pylyshyn, Z. (1986). Computation and Cognition: Towards foundation for cognitive science.

Quinlan, P. & Humphrey, N. (1987). Visual search for targets defined by combinations of color, shape and size: An examination of the task constraints on feature and conjunction searches. *Perception & Psychophysics*, *41*, 455-472..

Rao, R. P. & Ballard, D. H. (1995). An active vision architecture based on iconic representations. *Artificial Intelligence*, *78*, 461-505.

Rayner, K. & Fisher, D. L. (1987). Letter processing during eye fixations in visual search. *Perception & Psychophysics*, *42*(1), 87-100.

Reingold, E. M., Loschky, L. C., McConkie, G. W., & Stampe, D. M. (2003). Gaze-contingent multiresolutional displays: An integrative review. *Human Factors*, *45*(2), 307-328.

Rensink, R. (2000). The dynamic representation of scenes. *Visual Cognition*, *7*, 17-42.

Rensink, R., O'Regan, K., & Clark, J. (1997). To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science*, *8*(5), 368-373.

Robertson, G., Card, S., & Mackinlay, J. (1989). *The Cognitive Coprocessor Architechture for Interactive User Interfaces*. Paper presented at the 2nd annual ACM SIGGRAPH symposium of User Interface software and technology, New York, NY.

Roediger, H. L. (2003). Reconsidering implicit memory. In J. S. Bowers & C. D. Marsolek (Eds.), Rethinking implicit memory. Oxford, Uk: Oxford University Press.

Rogers, Y. (1999). What is different about interactive graphical representations? *Learning and Instruction*, *9*, 419-425.

Romani, G. L., Williamson, S. L., & Kaufman, L. (1982). Tonotopic organization of the human auditory cortex. *Science*, *216*, 1339-1340.

Rosch, E. (1973). Natural categories. *Cognitive Psychology*, *4*, 328-350.

Rosch, E. (1978). Principles of Categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and Categorization*. NJ: Lawrence Erlbaum.

Rosch, E. & Mervis, C. (1975). Family resemblance: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573-605.

Rosch, E., Simpson, C., & Miller, R. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 491-502.

Ryle, G. (1949). *The concept of mind*. New York: Barnes and Noble.

Sagi, D. & Julesz, B. (1985). "Where" and "what" in vision. *Science*, *290*(4704), 1217-1219.

Saunders, B. (2000). Revisiting basic color terms. *Journal of the Royal Anthropological Institute*, *6*, 81-99.

Scaife, M. & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, *45*, 185-213.

Schneider, G. E. (1969). Two visual systems. *Science*, *163*(3870), 895-902.

Schooler, J. W. & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, *22*, 36-71.

Schwartz, D. & Black, T. (1999). Inferences through imagined actions: knowing by simulated doing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *116-136*, 116-136.

Scialfa, C. T. & Joffe, K. M. (1998). Response times and eye movements in feature and conjunction search as a function of target eccentricity. *Perception & Psychophysics*, *60*(6), 1067-1082.

Shepard, R. (1962). Analysis of proximities: multidimensional scaling with an unknown distance function I & II. *Psychometrika*, *27*, 125-39, 219-246.

Shipp, S. (2004). The brain circuitry of attention. *Trends in Cognitive Sciences*, *8*(5), 223-230.

Shneiderman, B. (1981). *Direct Manipulation: A step beyond programming languages*. Paper presented at the Human Factors in Computing Systems CHI '86, New York.

Shneiderman, B. (1994). Dynamic queries for visual information seeking. *IEEE Software*, *11*(6), 70-77.

Shneiderman, B. (1996). *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization*. Paper presented at the In Proceedings IEEE Visual Languages, Washington, DC.
Shneiderman, B. (1997). *Direct manipulation for comprehensible predictable and controllable user interfaces*. Paper presented at the International Conference on Intelligent User Interfaces, Orlando, FL.
Shneiderman, B. (2002). Inventing discovery tools: combining informration with data mining. *Information Visualization*, *1*(1), 5-12.
Shneiderman, B., & Plaisant, C. (2006). *Strategies for evaluating information visualization tools: Multidimensional in-depth long-term case studies*. Proceedings from AVI workshop on BEyond time and errors: Novel evaluation methods for information visualization , New York.
Simons, D. & Rensink, R. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, *9*(1), 16-20.
Simons, D. & Silverman, M. (2004). Neural and behavioral measures of change detection. In L. M. Chalupa & J. S. Werner (Eds.), *The Visual Neurosciences*. Cambridge, MA: MIT Press.
Smith, B. A., Ho, J., Ark, W., & Zhai, S. (2000). Hand-eye coordination patterns in target selection. In A. T. Duchowski (Ed.), *Proceedings of the eye tracking research and applications symosium*. (pp. 117-122). New York: Association for computing machinery.
Sutherland, S. (1963). *Sketchpad, a man-machine graphical communication system*. Massachusetts Institute of Technology, Boston.
Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, *119*(2), 176-192.
Thibos, L. N. (1998). Acuity perimetry and the sampling theory of visual resolution. *Optometry and Vision Science*, *75*(6), 399-406.
Thompson, E., Palacios, A., & Varela, F. (1992). Ways of coloring: Comparative color vision as a case study for cognitive science. *Behavioral and Brain Sciences*, *15*, 1-26.
Todd, J. T. & Van Gelder, P. (1979). Implications of a transient-sustained dichotomy for the measurement of human performance. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(4), 625-638.
Torralba, A., Oliva, A., Castelhano, M. S., and Henderson, J. M. 2006. Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. Psychologial Review. 113, 4, 766-786.
Treisman, A. (1982). Perceptual grouping and attention in visual search for features and objects. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(2), 194-214.
Treisman, A. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97-136.
Treisman, A. & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*(1), 15-48.
Treisman, A. & Sato, S. (1990). Conjunction Search Revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(3), 459-478.
Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, Co: Graphichs Press.
Tufte, E. R. (1990). *Envisioning information*. Cheshire, Co: Graphics Press.
Tukey, J. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.
Tullis, T. (1997). Screen design. In M. Helander, T. K. Landauer, & P. Prabhu (Eds.), *Handbook of human-computer interaction*. (pp. 503-531). Amsterdam: North Holland.
Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327-352.
Ullman, S. (1984). Visual routines. *Cognition*, *18*, 97-159.
Ungerleider, L. & Mishkin, M. (1982). Two Cortical Visual Systems. In D. Ingle, R. Mansfield, & M. Goodale (Eds.), *The analysis of visual behaviour*. (pp. 549-586). Cambridge, MA: MIT Press.
van den Berg, R., Cornelissen, F. W., & Roerdink, J. B. T. M. (2008). Perceptual dependencies in information visualization assessed by complex visual search. *ACM Transactions on Applied Perception*, *4*(4), 22-41.
Vecera, S. & Farah, M. (1994). Does Visual Attention Select Objects or Locations? *Journal of Experimental Psychology: General*, *123*(2), 146-160.
Virsu, V. & Rovamo, J. (1979). Visual resolution, contrast sensitivity, and cortical magnification factor. *Experimental Brain Research*, *37*, 475-494.

Viviani, P. & Swensson, R. G. (1982). Saccadic eye movement to peripherally discriminated visual targets. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(1), 113-126.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. Journal of Experimental Psychology: Human Perception and Performance Journal of Experimental Psychology: General, 27(1), 92-114.

Wainer, H. & Thissen, D. (1988). Plotting in the modern world: Statistics packages and good graphics. *Chance*, *1*(1), 10-20.

Walker, R., Deubel, H., Schneider, W. X., & Findlay, J. M. (1997). Effect of remote distractors on saccade programming: Evidence for an extended fixation zone. *Journal of Neurophysiology*, *78*(2), 1108-1118.

Ware, C. (2000). *Information visualization: Perception for design*. San Francisco, CA: Academic Press.

Ware, C. & Mikaelian, H. H. (1987). *An evaluation of an eye tracker as a device for computer input*. Paper presented at the Proceedings of CHI + GI'87 Human Factors in Computing Systems.

Watson, B. A., Walker, N., Hodges, L. F., & Worden, A. (1997). Managing level of detail through peripheral degredation: Effects on search performance with head-mounted display. *ACM Transactions on Computer-Human Interaction*, *4*(4), 323-346.

Wickens, C. & Hollands, J. (1999). *Engineering psychology and human performance (3rd ed.)*. New Jersey: Prentice-Hall Inc.

Wickens, T. (2002). *Elementary signal detection theory*. New York: Oxford University Press.

Williams, D., Reingold, E., Moscovitch, M., & Behrmann, M. (1997). Patterns of eye movements during parallel and serial visual search tasks. *Canadian Journal of Experimental Psychology*, *51*(2), 151-164.

Williams, D. E. & Reingold, E. M. (2001). Preattentive guidance of eye movements during triple conjunction search tasks: The effect of feature discriminability and saccadic amplitude. *Psychonomic Bulletin & Review*, *8*(3), 476-488.

Williams, L. G. (1966). The effect of target specification on object fixated during visual search. *Perception & Psychophysics*, *1*, 315-318.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625-636.

Wilson, T. D. & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology*, *60*(2), 181-192.

Winaver, J., Witthoft, N., Frank, M., Wu, L., Wade, A., & Boroditsky, L. (2007). Russian blues reveal effect of language on color discrimination. PNAS, 104(19), 7780-7785.

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

Wolfe, J. M. (1998). What can 1,000,000 trials tell us about visual search? *Psychological Science*, *9*(1), 33-39.

Yantis, S. & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 601-621.

Yantis, S. & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 121-134.

Zacks, J., Levy, J., Tversky, B., & Schiano, D. (1998). Reading bar graphs: Effects of extraneous depth cues and graphical context. *Journal of Experimental Psychology: Applied*, *4*, 119-138.

Zadeh, L. (1965). Fuzzy Sets. *Information and Control*, *8*, 338-353.

Zadeh, L. (1982). A note on prototype theory and fuzzy sets. *Cognition*, *12*, 291-297.

Zelinsky, G. J. (1996). Using eye saccades to assess the selectivity of search movements. *Vision Research*, *36*(14), 2177-2187.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, *21*(2), 197-217.

Zhang, J. & Norman, D. A. (1995). A representational analysis of numeration systems. *Cognition*, *57*, 271-295.

Zhang, W. & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233-235.

APPENDIX A: English translation of Publication IV.

Published English abstract

Eye movements and the delusion of high acuity visual field

Our daily experience of vision seems to result from wide high acuity visual field. Nevertheless, we are unable to read without fixating to individual words. This apparent paradox has received attention during the past decade by the researchers of change blindness phenomenon. The subjects have been surprised at their poor ability to perceive changes. On the other hand, the research of experimental psychology has evidence for very fast abilities to make coarse judgments based on peripheral vision. Therefore, the experience of wide high acuity visual field is partly explained by overestimated belief in own abilities, and partly by unconscious constructive mechanisms that supplement the low resolution information from the environment. The identified mechanisms are 1) the coarse peripheral vision, 2) instantaneous information retrieval using eye movements, 3) the use of general knowledge for object recognition, 4) rapid conceptual memory, and 5) mental imagery based on longer interval memory.

Translated article

Introduction

Even important characteristics of the world go unnoticed by the people without their awareness. Attention is immediately drawn to rapid changes in visual field. However, simple and easily recognizable features might not be registered if some other factor interferes with this process. First, in the case of inattentional blindness (Mack & Rock, 1998), a demanding task prevents the person from detecting the presence of simple objects. For instance, majority of the subjects did not notice a lady with umbrella (Neisser, 1979) or a chest-beating person in a gorilla suit (Simons & Chabris, 1999) in the midst of a basket ball game if they had to count the number of passes.

Second, in the case of change blindness (Simons & Rensink, 2005), similar flaws of perception might occur in absence of any particular task. Even a large change in visual image might not be noticed. In the worst case, the whole image can be gradually changed and none of the steps is recognized. It has been presumed that the abrupt changes or movements of objects are required to capture the spatial attention so that they can be perceived (Rensink, O'Regan & Clark, 1997). However, the shift of spatial attention might be precluded due to eye blink, rapid flash of whole visual field, or visual changes in other object.

The perception might fail as a result from lack of attention or simply from inadequate visual resolution or acuity. Alternatively, the notion of visual resolution might include very many different kinds of neural representations and their interactions in human brain. In that case, inattention is simply a subclass insufficient resolution in the representation. The requirement for an accurate perception is that that the corresponding neural representation exists and is updated according to the current state of the environment. The source for inadequate resolution might as well be the low number of receptors in retina or the perceiver's inability to update the representation at the rate of changes in the environment. Thus, selective attention includes both focussed gaze and focussed information processing in perceptual system, termed covert attention. Traditionally, the explanations have over emphasized the latter. The underlying assumption has been that a person searching for an object is first focussing the processing of different spatial locations one at a time and only then directing the gaze towards the located target (Treisman & Gelade, 1980). The nature of covert attention is much debated and much more poorly defined as an explanation, and instead this article will focus on eye movements, peripheral vision, and transsaccadic memory. Findlay (2004) has argued that the covert attention is not really used independent of eye movements. The rapid parallel search from different location would then be founded on low spatial frequencies processed with peripheral vision.

The eye movements have been studied with different methods already for about one hundred years (Steinman, 2004 for review). The first measurements were based on the measurements of light reflecting from cornea and the fundamental challenge of this method is to differentiate between head and eye movements. Both are used for perception in normal circumstances. Therefore, the common method to restrict the head movements limits the generality of the findings. The first solutions to this challenge were the use of mirrors that were attached to the head. The longer measurement intervals were made possible by the use of mirrors attached to contact lenses. Eye movements have also been measured by the resulting magnetic fields. Due to the small angular size of high acuity vision, the eye movement measures provide useful information about which part of visual field is seen in detail and with colors. This spatial focusing is by far the most influential mechanism that controls attention.

This article will first discuss the findings about visual system in relation to peripheral vision. As the ability to discriminate fine details from peripheral vision is lower than from foveal vision, it is natural to ask what kind of perceptions or object-recognition it facilitates. Similar to peripheral vision also general knowledge of the perceiver has been shown to hasten the vision-based recognition. A patient suffering from Anton's syndrome might deny cortical blindness even though the primary visual cortex (V1) is destroyed. This represents an extreme case, where the recognition is entirely based on general knowledge. As an effect of general knowledge, it is interesting that the person does not recognize the source for the information for the misinterpretation and the constructed visual illusions are typical and believable given the prior experiences (e.g., Goldenberg, Müllbacher, & Nowak, 1995). Also the delusions of the high acuity vision discussed in this article refer to the inability to recognize the real source of information. In this case, the belief or sensation results from many constructive cognitive operations instead of light absorbed in the retinal receptors. Visual illusion is created if this cognitive construct differs from the visual information of the environment.

It has been suggested that the visual information is update on basis of demand. This notion is supported by the eye movement research of the everyday activities. On the other hand, the perceiver might experience having the high-acuity visual perceptual abilities, because they remember the visual percepts at earlier gaze directions. This persistence can be evaluated by the rapid serial visual presentation (RSVP) experiments. The alternative account is that like tropical fish humans would not remember high acuity information very long but would evaluate it only at the location of focussed central vision. For instance Horowitz and Wolfe (1998) performed RSVP experiments (with 111 ms presentations times) and generalized their results to an argument that the visual searcher would not learn anything at the longer intervals. In this article, learning and accumulation of knowledge across several saccades are examined by the research of visual imagery, as well. More precisely, how much the serial nature of collecting high acuity vision influences the later spatial knowledge in imagery? The discussed experiments attempt to find out whether the ability to later utilize the information perceived during one particular fixation is independent of other fixations. Alternatively, the ability to utilize detailed information can depend on the overview of the visual scene or context in general.

Influences of peripheral vision to perceptual abilities

The examples of inattentional blindness and change blindness mentioned in introduction are evidence for inability to perceive high-acuity visual information through the whole visual field. Unlike camera, human vision does not record unmodified detailed image and instead the ability to perceive details very accurately extends only to about 1° (degree of visual angle). When the distance from the center of vision (eccentricity) increases, the visual acuity deteriorates exponentially, and already farther than 4° (parafovea, Williams & Moody, 2004) visual acuity is much lower. The visual acuities at different locations of visual field have been evaluated by so called cortical magnification factor M that predicts the relationship between eccentricity and contrast discrimination threshold (Virsu & Rovamo, 1979). Later experiments with fMRI have supported the relationship with correlation between cortical area and discrimination performance. The lower discrimination thresholds at central vision was explained by the higher density of retinal cone cells, the increased sampling from cone cells by ganglion cells, and by the higher number of cells in thalamus and cortex that participate in the information processing.

Nevertheless, the peripheral vision serves an important role when the visual abilities are studied within the first few hundred milliseconds after opening of eyes. The perceiver does not have enough time to redirect gaze and high acuity vision for spatial locations relevant for the interpretations. In research literature, this state of visual perception has been called gist. Rensink (2000) has argued gist is used in a virtual representation process that creates the delusion of wide high-acuity visual field. On the level of performance, humans are only observed to be able to sequentially focus the gaze at different objects. This article does not assume that the mechanism described by Rensink would alone be sufficient to replace the lack of visual acuity in peripheral locations. The other relevant mechanisms utilize previous experiences. The perceiver both remembers detailed information from earlier gaze locations and applies the general knowledge about the environment.

Also others have proposed theories similar to Rensink's virtual representations (e.g., Simons & Silverman, 2004), but his theory was one of the first attempts to explain the delusion of high-acuity. According to Rensink's theory, the delusion is created by the sufficiently fast information access. Moreover, the perception is blurred before the focused foveal vision because the gist remains unchanged before the gaze is relocated. Rensink's notion of gist includes a schema that represents spatial relations between the objects. Alternatively, gaze direction is not required to change if it is sufficient that covert attention emphasizes the neural processing for the specified location. According to Simons and Silverman, perceptual experiences depend both on the assumed permanence of the environment and the finding that salient events exogenously attract attention and awareness. The event is simply not acknowledged if it fails to draw attention.

There are less controversial antecedent ideas in vision research for Rensink's theory. As the time goes by, the perception of a scene is thought to proceed from global to local elements. Navon (1977) specifies that details are perceived only if they are required by the interpretations. The delusion of high acuity follows from interpretations that are driven by conceptual knowledge and not only by the perceptions of the immediate environment. This hypothesis is supported by the finding that subjects claim to see human eyes in the facial photographs, even though they have been removed.

In the Navon's (1977) experiments, the subjects were shown letters (global elements) composed of even smaller letters (local elements). Subjects were able to willingly focus on either global or local elements. However, the global elements were recognized faster, and in the case of contradiction only the perception of local, and not global, elements were harmed. Navon's Stroop experiments (Stroop, 1935) provided similar results. The “ach” / “es” -sound discrimination was harmed by global H- and S-letters, but not similar local component letters. The results were corroborated by Kinchla and Wolfe (1979), who showed these findings extend up to the stimulus size of 6° - 9°. The largest stimuli used by Navon were the size 5.5°.

In order to describe the gist more accurately, its neural representations should be specified. This challenge is operationalized here as “What kind of questions can subjects accurately answer with gist?”. Biederman (1972) showed that the spatial relations of gist influence object recognition in a scene. The effect persisted even if subjects knew in advance exactly where the object would appear. The scene was divided into six sections that were scrambled (see figure). The part with to-be-recognized object was left unscrambled. Even though the scene was not large (3.5° x 5°), the scrambling harmed the object recognition with presentation times of 300, 500, and 700 ms. Thus, the subjects could have made few gaze shifts. Year later, Biederman (1973) showed that the findings generalize to visual search from large scenes (19°) as well.

The scrambling was especially harmful when the object was missing and its presence would have been likely given the surroundings. Furthermore, the target object was more often not found in the case of scrambling. The findings were interpreted as evidence for existence of scene level schema. Again, year later (Biederman, Rabinowitz, & Stacy, 1974), the phenomenon was studied and this time the effect of eye movements were controlled using limited presentation times (20, 50, 100, or 300 ms). With similar experimental paradigm, the effects of scene were observed even at the presentation times of 100 ms, and extending the presentation times to 300 ms did not improve the performance when the scene was scrambled. This time subjects were instructed to associate verbal labels to the scene. The performance improved when the task was to recognize the object from four alternative labels. To sum up, the results of Biederman's experiments show that peripheral vision influences the object recognition that requires high acuity. Short presentation time ensures that subjects were unable to make saccadic eye movements. Moreover, the function of peripheral vision is not merely to guide subsequent gaze shifts. With natural images, the influence of background is mediated by perceiver’s ability to associate meaning to visual perception and to evaluate the probabilities for presence or absence. The following section presents research findings that attach general knowledge to the associated meanings.

Figure The two stimulus stimuli similar to Biederman's (1972) experiments are presented above. Subject's task would have been to answer whether there is a fountain or not. In the lower picture, the segments were scrambled except the one including the fountain. The design of Biederman's stimuli did not control the discontinuity of the spatial frequencies at the border of segments. However, this difference alone is not likely to explain the results.

The influence of general knowledge and generalization

Thus, Biederman's experiments showed that the objects or persons (e.g., priest in a church) that were compatible with the background scene were recognized faster than the incompatible ones. The separate visual processes that specify the compatibility are associated with large (10°–15°) background features, diagnosticity of segregated objects (1°–3°), and the relations between these levels (Boyce, Pollatsek, and Rayner, 1989).

These experimental findings are supported by physiologically distinct systems of magno- and parvo-cellular visual pathways. The diagnosticity of an object defines its necessity for recognition process and its relation to target object by semantics or schema. Davenport and Potter (2004) showed that the compatible background increased the probability of accurate object recognition from 68 % to 82 % when the presentation time was shorter (80 ms) than the one required for saccadic eye movements. These results contested the earlier interpretations from experiments with line drawing images in which the object recognition performance was independent of semantics. Nevertheless, the influence of the surrounding scene is not limited to the compatibility of the schema.

Biederman (1982) named five relation factors that influence the recognition: (i) interposition of transparent objects, (ii) support of surfaces, (iii) probability of presence, (iv) probable position, and (v) familiar size. According to him, “the innocent bystander”-phenomenon proves that these factors cannot be explained by changing scene schema. Innocent bystander is an object, for which the recognition time does not increase when similar manipulations are done to other objects.

After the above-mentioned experiments, the defining criterion for scene related gist has been the performance that does not improve much when the presentation times are extended to allow saccadic eye movements. Chong and Treisman (2003) have used the same criterion to show that visual-perceptual representations utilize the information about group averages. In their experiments, the relatively good ability to evaluate the average of radii from a group of circles did not improve much (difference in threshold only 2%) when the presentation time was increased from 50 ms to 2 s. The ability to compare individual circles, however, increased (5 %). The ability to differentiate averages was almost independent of distributions of individual circles, with the exception that two same distributions are easier to compare than two different ones. These experiments were based on earlier experiments by Ariely (2001) with 500 ms presentation times. Ariely associated his findings with the experiments of visual search (Duncan & Humphreys, 1989), in which the reaction time depended on the ability to group features of visual objects according to their distributions. Also, Chong and Treisman associate averages and other statistical properties to distributed, as opposed to focussed, attention in visual searching. Visual-perceptual representations seem to be processing simple visual features according to statistical properties instead of qualities of individual objects. This is an obvious example of very fast perceptual constructivity. Thus, as one watches the numerous leaves of a tree, the perception represents some kind of a generalization instead of features of individual leaves.

The better representation of an average compared to individuals can be interpreted as a more accurately defined phenomenon than the ones Rosch (1978) characterized in relation to prototype theory. Rosch's earlier observations emphasized the importance of typicality for perception (Rosch, Simpson & Miller, 1976). Prototype theory states that percept is not compared to typical or atypical individuals, but a summarizing and generalizing representation of typical features. On top of that, Chong and Treisman speculated that representations of averages could be useful for detecting outliers and differentiating objects from textures. In relation to gist, these functions are essential when the proper location for next gaze direction is evaluated.

The influence of rapid memories

In the above-discussed experiments, the meaning in relation to previous knowledge about the environment and its statistical properties influenced the perception. The memory experiments discussed in the following are extending the role of semantics. Potter (1976) showed that pictures can be understood at faster rates than they are actually memorized. In these RSVP experiments, the presentation time for one image was 113 ms. Even so, subjects could determine in 60 % of the cases whether image corresponded with verbal description. Later memory experiments (Potter, 1993) however showed that, after the sequence, subjects were unable to determine whether they had seen the image, or not. The encoding of a picture into conceptual memory is interfered by the following image of a sequence up until 1000 ms. According to Potter, conceptual interference requires attentional resources unlike perceptual interference. The formation of an enduring memory in the experiments required even 400 ms of non-disrupted time. Potter (1993) interpreted the results of her later experiments in a way that the subjects were able to construct their own concepts (e.g., picnic) and not just select the predetermined categories. This ability was tested by semantics-based decisions from alternative words that were not presented until the recall phase. The conclusions of Biederman's (1982) experiments were similar; some semantic relations between the objects were interpreted so quickly that they interfered with recognition of these objects.

Even though iconic memory can persist up to 300 ms (Irwin, 1996), 400 ms were needed for encoding of memories (Potter, 1976). The serially presented pictures primarily influence each other as interruption in encoding of memories. The memorizing of several images is disturbed only at very general level of processing and not on the level of individual pictures (Potter, 2002). The ability to perceive a change can also be prevented by a flash following immediately after the image. Rensink, O'Regan and Clark (1997) used this experimental procedure to show that the change is not perceived even though the original and changed image were shown repeatedly one after another in intervals of 240 ms if empty screen was flashed in between for 80 ms. Subjects knew that some kind of a change is going to happen. Additional important finding is the persistence of this phenomenon when presentation times are extended to include eye movements. Without the flash, 85 % of the disappearing objects were detected by peripheral vision in other experiments (Parker, 1978).

Furthermore, the subjective importance of the object influences the ability to perceive changes in it. Thus, also this factor points to the significance of attention in explanations of perceptual abilities. To sum up, change blindness measures how much details the memory includes. Potter's experiments, on the other hand, show that the memory is rapidly fading if the temporally activated perceptual information is not attached to the structure of semantics or schema.

Directing gaze in everyday situations

Semantics also influence the process of gathering new information. The research of eye movements primarily reveals the locations from which user hopes to gain new information. Inversely, this sometimes allows the researcher to evaluate what subject already knows. In general, the gaze is not so much directed to obvious items, but to surprising ones. The requirements of high acuity vision and selective focus of gaze, however, depend on task. In some experiments (Land, 2004), the perception of high-resolution image was not deteriorated when the visual field was dynamically limited to 10° by using gaze-contingent method (in which the limiting frame moves along with the gaze). Therefore, the information from peripheral vision outside 10° was not acknowledged to create the delusion of high acuity visual field. Before evaluating the role of eye movements in creating the delusion any further, some aspects of the physiological foundations of visual processes need to clarified.

To begin with, recorded eye movements are generally categorized according to their speed and degree of voluntary control. The fast movements are called saccades. According to the research traditions, slower movements with corresponding objects are called smooth pursuit and the whole-field-wide movements are called optokinetic reflex. On the level of behavior, however, these classes form continuum. Head movements, and not only eye movements, control the gaze direction as well. This is most apparent in the case of vestibular-ocular reflex, when moving the eye into opposite direction compensates the rotation of head. The process utilizes the sensory information of vestibular organ in inner ear. For the purposes of this article, it is interesting to note that the requirements of compensation depend on the gaze point. The required compensation is smaller when the eyes are fixated object at long distance. These compensation processes relate to a more general characteristic of the visual system. It has developed almost exclusively for detecting changes in retinal image that are caused by eye movements. If the retinal stabilized image – for instance at contact lens – does not change, the resulting neural representation disappears very quickly. Thus, an efficient neural processing does not represent the unchanged aspects of the environment, and this applies to the delusion of high acuity vision as well (Simons and Silverman, 2004). In a nutshell, the information is acquired only when needed.

The following discussion will focus on how the eye saccades are utilized to collect information about the environment. The transsaccadic integration of retinal image, before and after the saccade will not be discussed much further. However, the amount of information accumulated is surprisingly small and the attention related limitations (so called capacity) prevent one from maintaining the location information for very many (over four) objects (Irwin, 1996). The importance of selecting saccadic targets is emphasized by the findings that more details are remembered from close proximity of target location. The targets of high acuity vision are of interest when the natural and ecologically valid vision is researched.

Already in 1960s, it was known that saccadic targets depend on objectives of the perceiver (Yarbus, 1967). During normal kitchen work (e.g. making tea) gaze is almost exclusively directed to objects, one at a time (Hayhoe et al., 2003; Land, 2004). First the whole body is oriented towards the object, next moves the eyes, and finally the hands extend for grasping. It seems that the high acuity vision supports the operations with hands, but not the general orienting with the body. If the person is instructed to maintain a constant fixation point, the required time is tripled. Only 5% of the time task-irrelevant objects were gazed, and therefore endogenously driven spatial attention deserves to be discussed later in this article. The acting person also makes use of the ability to predict events when locating the gaze. Drivers gaze might already adapt to changes before the steering wheel is rotated (Land, 2004). Nevertheless, inexperienced drivers alone use high acuity vision for holding the car on a lane. Likewise, the gaze of a table tennis player who is receiving the serve precedes the bounce from the table with 200 milliseconds.

When the gaze location is almost constant, it is called fixation. According to Land (2004), the purpose of fixation is to locate something, participate in the grasping, or to confirm some condition. However, the hand itself was not gazed in his experiments. When reading the average duration of fixations is 225 ms. Furthermore, the masking experiments indicate that only the first 50 – 70 ms are used for receiving visual information (Land, 2004). Thus, most of the time is spend on preparing the next saccadic eye movement. For natural images, the average fixation duration is 330 ms, but there is relevant variance between individuals (Henderson, 2003). In every day tasks, such as when preparing a sandwich, the fixation duration varies even between 100 – 1500 ms (Hayhoe et al., 2003). The longest fixation durations are associated with the guidance of grasping and the shortest are apart of a predefined sequence. These findings emphasize the process nature of directing gaze as a part of visual perception, similar to Rensink's (2000) virtual representations. The gaze direction in context of mental imagery is discussed. The aim is to determine whether this process nature is only specific feature of information pick up, or is the neural representations constructed sequentially, as well.

The eye movements during mental imagery and cumulative visual memory

The first indications of the role of eye movements in mental imagery were reported in 1960s. Haber and Haber (1964) studied the eidetic memory of 8 – 12-year-old children. The eidetic memory roughly resembles the popular conceptions of so called photographic memory. The following qualitative differences to others were identified with about 8 percent of children in this age group, but almost none with the older age-groups. Their memory (i) was not based on after images, because it lasted dozens of seconds (eye movements are within hundreds of milliseconds) and (ii) included positive colors. They remembered (iii) many small details and (iv) experienced the images as lively and external to them. In addition to verbal descriptions, the eidetic memory was distinguishable in (v) eye movements that corresponded with spatial location of the previously shown details. Hebb (1968) was, to my knowledge, the first who interpreted these findings so that eye movements have the role of organizing memories.

The interpretation was ahead of its time and did not become empirically meaningful until when the measurement devices developed the same level as in the experiments discussed in the following. In the 10 years, Brandt and Stark (1997) were the first to report controlled experiments about the association between the content of mental imagery and eye movements. The significant factor that made their experiments successful was their choice of the task and stimuli. The task was to remember the location of small quantities of black dots within a chessboard. Thus, all the meaningful content of the experiment explicit in the visible spatial organization. The subjects' eye movements at the moment of perceptual encoding and at the moment of mental imagery strongly resembled one another. However, the fixations duration was 20 ms longer in the case of mental imagery and, as a result, the total gaze path was shorter within the equally long interval. Process monitoring hypothesis explains the lengthening of fixation by the difficulty of constructing the memory. The similarity of gaze paths was inferred by using string-editing analysis; every location is marked with specific symbol, and then, the similarity of strings of these symbols is analysed. It is not clear whether these findings hold if the semantic significances of the original locations are changed. To my knowledge, this is yet to be tested.

Laeng and Teodorescu (2002, see also Mast & Kosslyn, 2002) used the experimental setting of Brandt and Stark to prove the causality of eye movements in memory-based imagery. The resulting eye movements, in the experiments of Brandt and Stark, could have only been secondary and epiphenomenal effect of imagery. In the experiments of Laeng and Teodorescu, there were no eye movements if the eyes were fixated when subjects perceptually encoded the image. Note that it is possible to redirect attention without moving eyes. This is called covert attention. However, the performance deteriorated if it was possible to move the eyes during the perceptual phase, but not during imagery phase. Thus, the locations are represented in a coordinate system relative to gaze point. This finding was confirmed also by another experiment. In that experiment, subject used visual imagery to answer a question about visual details of a tropical fish. Fishes were placed into each of the four corners of the image. The authors' interpretation of the results was that the mental image is constructed using spatial index of a motor coordinate system. This is how the order of parts is encoded. Thus, perception is not only about storing the representations, but there is also learning about directing attention and searching or browsing the objects. Furthermore, Chapman (2005) compared the memorizing of fixated and non-fixated objects. According to his results, memories of pictorial details depend on gaze locations, in contrast to memories of more general relations that are part of the gist.

A similar new interpretation of change blindness has emerged in recent years. According to Henderson and Castelhano (2005), change blindness results from shifting gaze instead of covert attention. Their experiments controlled the eye movements. An object was masked only after it had been fixated and the gaze had shifted to another location. Consequently, the number of change detections increased above the random level. The difference to the results of Rensink et al. (1995) ensue from making sure that the subjects had been fixated the later disappearing object. Several subjects also acknowledged the change after fixating again the changed location that acted as a memory cue. Thus, Henderson and Castelhano argue that change blindness often result from problems of recall that are related to the current gaze point. To summarize, it appears that, at least part of, visual memories are represented in a coordinate system of eye movements. If the representations simply utilized the coordinate system of external world, the eye movements would be unnecessary and shift of covert attention would be sufficient.

Summary

This article has discussed the cognitive processes that are responsible of constructing visual representations. Navon's (1977) experiments imply that peripheral vision is first used to construct a representation of larger regions as part of a gist and only then spatial attention selects some details of the scene. According to Navon, relevant details are located before they are identified. The first line of research discussed in this article examined the information perceived during one fixation. The effects of eye movements were excluded by short presentation times. The results of Biederman et al. (1972; 1973; 1974) have shown the importance of peripheral vision in object recognition. He interpreted the object recognition to depend on both more general scene schema and spatial relations between the objects in the scene. Thus, perceiver first creates gist level representation, which is used for more precise representations. During a fixation, general knowledge is used to interpret the relation between the object and the scene. For instance, based on the earlier experiences, priest is easier to recognize at the church. At the same time, new relational knowledge is generalized. Ariely (2001), and researchers Chong and Treisman (2003) have shown that averages are extracted quickly and even more reliably than feature values of individual object forming a group. Therefore, rapid perceptual representations mostly involve statistical properties of groups, instead of individuals, in the same way that gist precedes details. In the experiments of Biederman et al. (1982) and Potter et al. (1976; 1993; 2002), semantic relations between objects were rapidly available to the perceiver and influence object recognition. Thus, high acuity vision depends on understanding semantic relations of the details.

In addition to generalizations, gist can be utilized for conceptualization. Potter's experiments showed that scene can be conceptualized as picnic very rapidly, but it is still not recognized as familiar soon after. Recognition memory requires time without interruptions for encoding. Similar ideas have been used to explain change blindness. According to this explanation, brief flash can prevent encoding the changing detail. The assumed requirements, that are not fulfilled, are the amount of details in the case of change blindness and the association to conceptual knowledge in the case of Potter's RSVP experiments.

Rensink, as an important change blindness researcher, has argued that this phenomenon proves the lack of high acuity vision at the corresponding retinal location. According to him, the delusion of high acuity in vision results from knowledge that is accessible immediately - from the perspective of experience – by directing the gaze. As a metaphor, he describes an Internet page that is accessible by hypertext links, as opposed to having the information stored at own home computer. Consistently, gaze directions in every -day tasks often focus on objects at hand. Even if the peripheral gist would be enough for the task, the motor performance time triples when the eyes are not moved. Accumulated knowledge is used to anticipate the future needs for high acuity vision. The findings from more controlled experiments suggest that very much information is not accumulated over individual saccadic eye movements. Furthermore, at least some of the accumulated memories are not represented in one coordinate system independent of gaze-point. The exact processes to accumulate information across gaze points are currently active subjects of research.

In addition, the research of visual mental imagery has found support for gaze dependent memory representations. The research of school children in 1960s provided indications that efficient memory performance requires repeating the original eye movements. The controlled eye movement research methodology has confirmed these observations about ten years ago. Later experiments have also confirmed the causal role of eye movements for quality of mental imagery. If eye movements were prevented, memory performance was harmed. Thus, spatial memory seems to be represented in brain relative to gaze point of original perception. Again, similar explanations have been used to explain change blindness phenomenon. According to Henderson and Castelhano, change blindness is often caused by problems in retrieving the information from memory. The performance improves if there is certainty that the subject had fixated at the location of change before and after the change.

Discussion

This article has evaluated the quality of information that is available for the perceiver from the period of one eye fixation or from longer intervals. The reviewed experiments have supported the argument that all the high acuity visual information is not received at once, and instead, the perception is influenced by subjective construction of representations. The following components of the construction process were identified: the general knowledge from consistencies of the environment, short-term accumulation of memories and low acuity peripheral vision. The researchers do not have direct access to these accumulated representations, and their nature is inferred from behavior in controlled experiments. For instance, the findings about change blindness do not directly prove that there is no high acuity representation of the whole visual field. However, the capacities that are not utilized are outside of the scope of empirical behavioral research. Thus, this article has focused on the positive findings; the capabilities in special circumstances (e.g., rapid presentation times) and the tendencies in every-day situations.

In the event of change blindness, it is possible that the perceiver do not simply compare the representations before and after the change (Simon & Silverman, 2004). The explanation of Henderson and Castelhano is a well-specified process model that exemplifies this possibility. The blindness that results from inattention can also be explained by inability to consciously acknowledge the arrival of an object that is in fact perceived. For instance, if these subjects are forced to choose between this and few other alternatives, they select the correct alternative more often (36 %) than incorrect one (4 %) (Mack & Rock, 1998). The rate of correct choices is relatively high, because even the subjects who were consciously aware of the object arriving were accurate about the identity only in half of the cases.

Change blindness is often not important or acknowledge during every-day activities because the gaze directing is used for receiving the necessary information. Especially during the motor activities, gaze is fixated to an object only as long as the required information is obtained and the immediate purpose of each fixation is independent of the other fixations (Hayhoe et al., 2003). Accordingly, Hayhoe and her colleagues have interpreted their findings as a support for sensorimotor theory, which argues that the visual scene acts as external memory. Perception has a special function during actions. Nevertheless, the ability perceive changes is also influenced by visual saliency of the features: size (spatial frequency), density of edges, local contrast differences, and correlation between adjacent points (Henderson, 2003). However, when the scene is meaningful, the effect of saliency diminishes. Subjective interests have significant influence on selecting the gaze direction (Yarbus, 1967). The prerequisite for this effect is the gist-related information about the spatial regions of interest (Navon, 1977). Alternatively, studied subjects have been shown to make scanning eye movements as a preparation for the following actions (Hayhoe et al., 2003). The high acuity vision is also likely to supplement gist-related knowledge and object identities in the natural circumstances.

Horowitz and Wolfe (1998) have even made the exaggerated suggestion that there would no accumulation of information at all between the fixations. In their experiments of visual search, the subjects did not learn by making use of the unchanged objects. Peterson et al. (2001), however, pointed out that these subjects had learned to become faster even though the general accuracy of decision had not improved, when the object were not changed. The amnesia interpreted by Horowitch and Wolfe is likely to result from the presentation times that were shorter than the latency for eye movements. Peterson et al. increased the presentation time and force the search to become serial (include several eye movements). As a result, subjects still re-fixated the unchanged objects, but not so often than when the objects were changed. The idea of accumulation is also supported by Chun's (2000) theory of contextual cueing. In his experiments the decisions about already seen natural images were more accurate than with completely novel images.

The more philosophical literature has explained the purpose of visual perception in two ways. First explanation argues that perceptual system attempts to represent the external world as accurately as possible (especially in Marr, 1982). The alternative sensorimotor theory (O'Regan & Noe, 2001), however, is currently more popular and argues that the task of perception is to facilitation of actions (Gibson, 1972; 1977, more lately Noë & Thompson, 2002). Sensorimotor theory emphasizes the perception of contingencies that result from interaction of environment and actions. Similarly Neisser's (1976) theory of perceptual cycle strives toward more interactive perspectives of perception. In the theory, the information is picked up, compared to expectations and used to guide the actions, such as gaze direction. If the received information is consistent with the expectations, the focusing fixation is not needed in that location. Accordingly, the experimental results show that the gaze is fixated in average one fixation earlier to the changed objects compared to unchanged ones (Parker, 1978). Thus, peripheral vision influences the later accumulation of visual information.

Aalto-DD 30/2012

ISBN 978-952-60-4548-1
ISBN 978-952-60-4549-8 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934
ISSN 1799-4942 (pdf)

Aalto University
School of Science
Department of Media Technology
www.aalto.fi

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

DOCTORAL
DISSERTATIONS