

Tilburg University

Data for all

Quispel, Annemarie

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Quispel, A. (2016). *Data for all: How professionals and non-professionals in design use and evaluate information visualizations.* [s.n.].

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Data for all

How professionals and non-professionals in design use and evaluate information visualizations

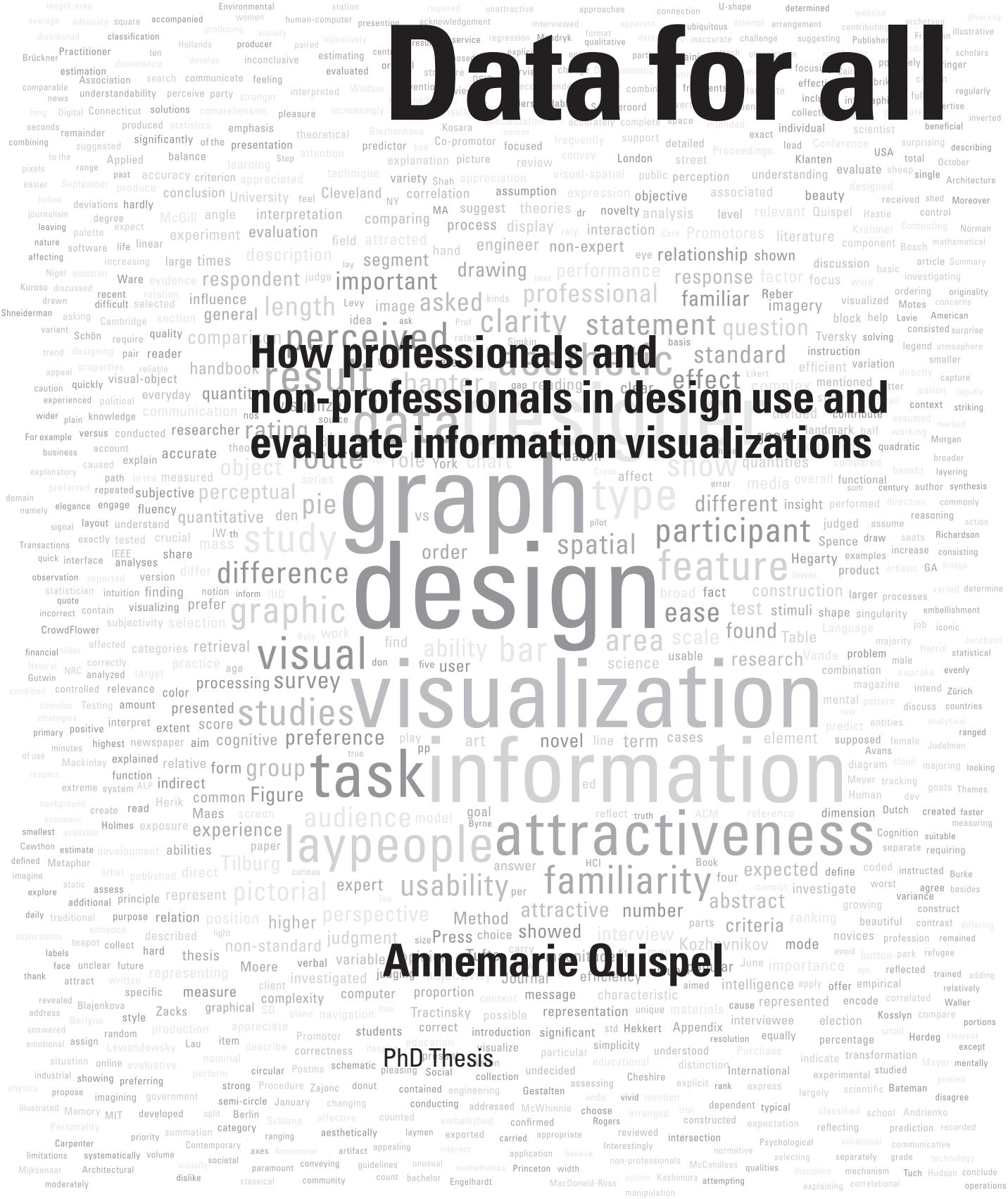
graph design visualization

task information visualization

people attractiveness

Annemarie Quispel

PhD Thesis



To my mother and my father

Data for all

**How professionals and non-professionals in design
use and evaluate information visualizations**

Colofon

Data for all. How professionals and non-professionals in design use and evaluate information visualizations

Annemarie Quispel
PhD Thesis
Tilburg University / Avans University of Applied Sciences, 2016

TiCC PhD Series No. 45

The research in this thesis was conducted with financial support of Avans University of Applied Sciences and the Centre of Expertise Art & Design of Avans University.

isbn/ean: 978-90-9029761-3
Design and layout: Koen van der Weide
Print: MK Publishing

© 2016 Annemarie Quispel

No part of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without written permission of the author or, when appropriate, of the publishers of the publication.

Data for all

**How professionals and non-professionals in design use
and evaluate information visualizations**

Proefschrift ter verkrijging van de graad van doctor
aan Tilburg University
op gezag van de rector magnificus,
prof. dr. E.H.L. Aarts,
in het openbaar te verdedigen ten overstaan van een
door het college voor promoties aangewezen commissie
in de aula van de Universiteit
op woensdag 15 juni 2016 om 14.15 uur

door Annemarie Quispel
geboren op 19 september 1967 te Schiedam

Promotor

prof. dr. A.A. Maes

Copromotor

dr. J. Schilperoord

Promotiecommissie

Prof. dr. E.J. Kraahmer

Prof. dr. E.O. Postma

Prof. dr. J.A.L. Hoeken

Dr. M.F. van Dartel

Dr. L. van Weelden

Content

Chapter 1	9
General introduction	
Chapter 2	21
Information visualization for a general audience: the designer's perspective	
Chapter 3	43
Would you prefer pie or cupcakes? Preferences for data visualization designs of professionals and laypeople in graphic design	
Chapter 4	63
Graph and chart aesthetics for experts and laypeople in design: The role of familiarity and perceived ease of use	
Chapter 5	85
Reading graphs. The role of length and area in comparing quantities	
Chapter 6	103
Visual Ability in Navigation Communication	
Chapter 7	117
General conclusion and discussion	
References	127
Summary	137
Acknowledgements	140
TiCC Ph.D. Series	141

Chapter 1

General introduction

1.1 Introduction

Information visualizations are traditionally used by scientists and other professionals in analytical tasks. But they are increasingly used in mass media, not with the purpose of analyzing large numbers of data, but to inform a broad audience of non-experts about facts and trends in society. See for example the graphs below.

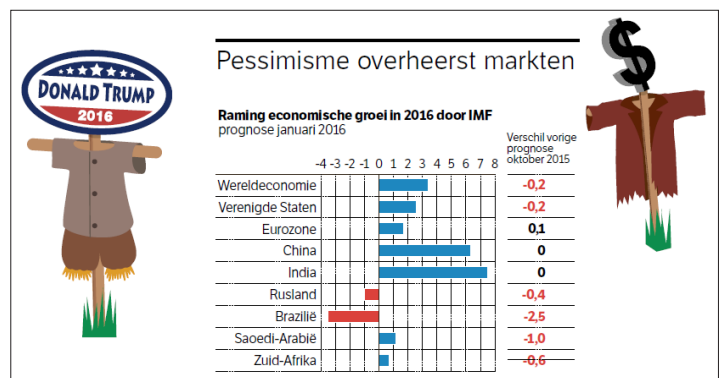
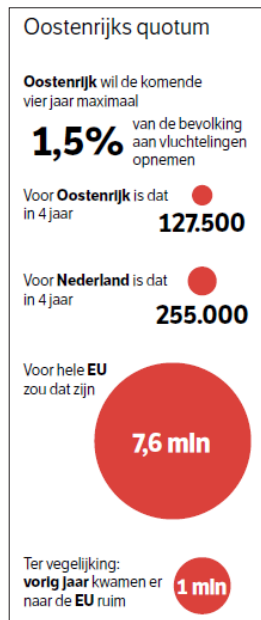


Figure 1 Proposed refugee quota, NRC, January 2016

Figure 2 Economic growth estimate, NRC, January 2016

Figure 1 shows quota proposed by Austria for allowing refugees, and the numbers of refugees this would result in in Austria, the Netherlands, and in the whole EU, the latter also compared to a year before. Figure 2 shows pessimistic estimations of economic growth worldwide. What is different in these examples from traditional ways of visualizing?

First, 'popular' information visualizations increasingly use novel ways to visualize quantitative information, novel ways to 'encode' abstract data into a graphical form. The most familiar way to represent quantities is in the form of bars in a bar graph (as in Figure 2). But in the refugee example, quantities are represented in the form of 'bubbles'. The sizes of the bubbles express the sizes of the quantities. Other examples of novel types of information visualizations that are sometimes found in mass media are 'donuts' and semi-circles. In these designs quantity is represented by segments of a circular or semi-circular bar.

Second, information visualizations in mass media are not always plain and abstract, as in science and statistics, but are often 'embellished' with pictorial elements. In Figure 2, lack of economic growth is visualized in a traditional bar graph, but additional information is given by pictorial elements in the form of

scarecrows, illustrating reasons why financial markets are pessimistic. They may embellish graphics, but can at the same time produce some confusion, for example because it is unclear in this case whether the position of the scarecrow (left or right from the graph) is relevant or not. In the example below (Figure 3), the use of a novel visualization technique and pictorial elements have been combined.

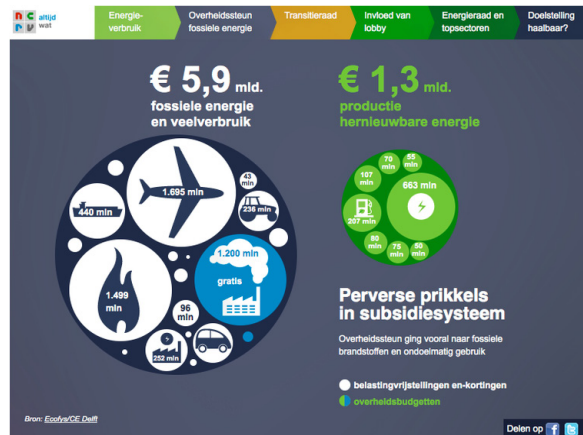


Figure 3 Subsidies for energy in the Netherlands, by Karin Schwandt (for NCRV Dutch tv broadcasting)

This example shows amounts of government subsidy for various types of energy. Quantities are represented by the sizes of the bubbles, and the energy categories are not explained by verbal labels, but by pictorial icons such as an airplane, car, or factory.

Thus far, the study of Information visualization mostly focused on visualizations allowing an accurate and efficient reading of data. Numerous studies have investigated features that enhance their effectiveness. Far less is known about what makes a 'good' information visualization for a broad audience. What criteria do designers use for such visualizations? To what extent do they consider adequacy, understandability, and attractiveness important? And what is the effect of using novel visualization techniques and pictorial elements on their understandability and attractiveness? Similarly, little is known about the way the general public understands and appreciates these visualizations. To what extent do they share opinions with the designers about the importance of clarity and attractiveness, and about what makes a visualization attractive?

This is what this thesis is about: information visualizations for a broad audience: how are they produced, understood, and evaluated by their producers, design experts, and by their audience, laypeople in design? What are the main criteria, and (how) do these criteria differ for designers and laypeople? The visualizations we focus on differ in degrees of abstractness, with a main focus on the visualization of abstract graphs, visualizing quantities. In the remainder of this chapter we first discuss the societal and theoretical relevance of the thesis. Subsequently, we describe the methods we used and introduce the studies that are described in the remaining chapters.

1.2 Societal relevance

In this thesis, we address four research questions regarding data visualizations for a broad audience:

- 1 What is the importance of functional and aesthetic criteria in judging visualizations?
- 2 What makes popular information visualizations attractive?
- 3 What makes information visualizations usable?
- 4 How do designers and laypeople differ in their understanding and aesthetic preferences?

In this section we discuss the societal relevance of this research.

Investigating information visualization is relevant for a number of reasons, which are discussed more elaborately below: enormous amounts of data need to be visualized for the general public; there is a lack of knowledge about the way 'popular' visualizations are understood and appreciated; information design and designers are increasingly important, but little is known about design practice. Gaining more insight into information visualization would be beneficial for design education and practice, and, eventually, the general public.

People are facing massive amounts of information every day. Architect and graphic designer Richard Saul Wurman (2012) states that much of this information concerns raw data that somehow need to be transformed to become meaningful information. Data have become widely available, thanks to rapid developments in information technology, but also thanks to journalists and bloggers demanding freedom of data, and to governments striving for transparency, as data journalist Simon Rogers of The Guardian describes (2012). For example, Barack Obama opened a portal for government data in 2009, offering public access to over 188.989 data sets (www.data.gov) about business, education, climate, health, et cetera. This initiative has been followed by several other countries, including the UK. For example, the national newspaper The Guardian offers the full datasets behind its news stories, which attract a million page impressions a month.

Much of the data that we are bombarded with can best be understood by visualizing them (Yau, 2011). This is being done by an increasing number of designers, including many experts in graphic design. The term 'graphic design' refers to both the act and the final product of conceiving, planning, selecting, organizing and shaping a series of elements – usually a combination of text and images – for the creation and presentation of visual communication (Frascara, 2004). The term 'graphic' in graphic design refers to the printing techniques used to produce and distribute products such as books, magazines, newspapers and posters, but graphic design also encompasses a wide range of activities and products typical for the 'digital age', like the design of websites,

apps, and information visualization. We see a growing number of information visualizations being published in mass media. We also see a growing variation of such visualizations. Designers, whose job is the ‘conception and realization of new things’ (Cross, 1982), do not confine themselves to conventional visualization techniques (e.g. bar and pie graphs), but develop novel ways of visualizing information (as in Figure 3). The question then arises to what extent these novel types of information visualizations are understood and appreciated by their audience of laypeople. What makes them effective for everyday tasks to be performed by a broad, non-expert audience, such as assessing which political party has won the elections, or judging how many more refugees are going to be allowed in the EU compared to a year before, as in Figure 1? Gaining insight into the way these visualizations are understood and appreciated by their audience would be beneficial for designers and, eventually, for their audience.

Little is known about designers’ ways of working. Designers have a great deal of responsibility in the way information is visualized to inform a general audience about, and to engage them in developments that affect their life and society. Moreover, design has become a significant economic sector. According to the Dutch central bureau of statistics (CBS) there are about 47.000 registered designers in the Netherlands in 2007, about half of whom received design education, mostly in graphic design. Unlike scientists, graphic designers are not used to document their ways of working. The graphic design field lacks a self-definition that can support and integrate research (Storkerson, 2006). Further, designers are used to work on the basis of intuition and experience, rather than explicit knowledge (Polanyi, 1966; Cross, 1982; Schön, 1983). Designers, just like most other professional practitioners, are not used to explicitly document their methods and professional practice. As Friedman (2003) states, designers could benefit from the insights that studies into the graphic design practice can provide, as these could enable them to move from solving one unique case after another to broader explanatory principles and solutions for similar kinds of problems.

This thesis contributes to a better understanding of the designers’ practices, of the quality criteria used by designers and their audiences, and of design characteristics determining the usability and attractiveness of such information visualizations.

1.3 Theoretical relevance

In investigating information visualizations for a broad audience, this thesis takes two perspectives: the perspective of the designers, i.e., the producers of visualizations, and the perspective of the laypeople, the non-expert users of popular information visualizations. The thesis investigates what the main

criteria are for 'good' information visualizations, what factors contribute to the usability and attractiveness of information visualizations, and how traditional and novel forms of visualization are understood and appreciated by designers and laypeople. The studies in this thesis are confined to static 2D data visualizations as they are published in printed mass media, representing a combination of nominal and quantitative information (like in a graph visualizing election results).

In the remainder of this section, we discuss the theoretical relevance of the main questions addressed in the thesis.

1 *What is the importance of functional and aesthetic criteria in judging visualizations?*

Most empirical research into information visualizations, in particular the study of graph design and comprehension, has focused on clarity: the accuracy and efficiency with which specific tasks can be performed with them (e.g. Mackinlay, 1986; Scaife & Rogers, 1996; Carpenter & Shah, 1998; Kosslyn, 1994; Ware, 2004, 2008). Far less is known about what makes a good information visualization for a broad audience. In the scientific field of information visualization, aesthetics has come to be recognized as an important research subject (e.g. Chen, 2005; Burkhard, Andrienko, & Andrienko, 2007). Aesthetics is a complex, multifaceted notion, and the term is used in various meanings in different realms, ranging from appreciation of art works and beauty to pleasing the senses. In this thesis, as in many of the studies we refer to, with the term aesthetics we refer to attractiveness.

Researchers in information visualization assume that aesthetics is important in engaging a lay audience in the information (e.g. Kosara, 2007; Vande Moere & Purchase, 2011), and advocate cooperation between scientists and designers to reach an optimum synthesis between usability and aesthetics (e.g. Kosara, 2007; Judelman, 2004; Vande Moere, 2005). These researchers assume that designers put much emphasis on aesthetics, and aim to express subjective concern, rather than communicating objective information. As Vande Moere and Purchase (2011) put it: '(...) complex and socially relevant issues might best be communicated to a large audience through popular media using an artistic and engaging visualization (even if its designer knows that such a method is not the most effective or efficient).' (p.361). According to Kosara (2007) 'artistic' and efficient forms of visualizations seem irreconcilable: 'Visual efficiency does not play a role in artistic visualization, quite the contrary. The goal is not to enable the user to read the data, but to understand the basic concerns.' (p.634). But little is known about the criteria designers themselves use when they design data visualizations for a broad audience, and how they think about the relative importance of usability and aesthetics. Designers mostly work on the basis of experience, intuition, and trial and error (Schön, 1983). Similarly, little is known about the way their audience of laypeople weigh functional and aesthetic criteria.

2 *What makes popular information visualizations attractive?*

Despite an increasing acknowledgement of the importance of aesthetics in information visualization, little is known about what makes a visualization attractive. In this thesis we investigate the effect on attractiveness of three features: the use of pictorial elements, novelty, and clarity.

The use of pictorial elements is one of the characteristics associated with 'popular' information visualizations. Some designers, such as Nigel Holmes, believe that graphs embellished with pictorial elements are found attractive by a broad audience of non-expert readers (Holmes, 2006). Others, such as Edward Tufte, believe that 'graphical elegance is in simplicity of design and complexity of data' (Tufte, 2001, p.177), thus rejecting all sorts of embellishment. Some empirical studies have addressed the influence of using pictorial elements on preferences for graphs (e.g. Bateman, Mandryk, Gutwin, Genest, McDine, and Brooks, 2010; Tractinsky & Meyer, 1999; Levy, Zacks, Tversky, & Schiano, 1996). But these studies did not directly address the question which of the two graph types – abstract or pictorial – is judged most attractive. The focus on pictorial elements is also relevant for another reason. Recent psychological studies have shown differences in visual intelligence between designers and non-designers (e.g. Blazhenkova & Kozhevnikov, 2010), which cause differences in attention to, amongst other things, pictorial details vs. schematic spatial relations. This might lead to differences in aesthetic experiences, particularly for pictorial vs. abstract visualizations.

Another characteristic often associated with popular information visualizations is the notion of novelty. Designers are said to often use more or less 'novel' ways to represent quantities, with novelty being considered the counterpart of familiarity. Aesthetic theories often use these or comparable terms to explain why and when visualizations are attractive. According to the theory of evolutionary aesthetics, human beings derive aesthetic pleasure from phenomena that help them survive (e.g. Hekkert, 2006). On the one hand they are attracted to familiar things, because familiarity, as a result of repeated exposure, facilitates perceptual organization and helps them to bring order in a complex world. On the other hand people are also attracted to new, unusual things, presumably because novelty facilitates learning, which is also a vital capacity for survival. Other theories predict that attractiveness increases with increasing familiarity (e.g. Zajonc, 1968, 1984; Reber, Schwartz, & Winkielmann, 2004). According to processing fluency theory, repeated exposure to a stimulus results in familiarity, which in turn makes perceptual and cognitive processes more fluent, and this fluency is perceived as attractive (Reber et al., 2004).

In view of the growing number of novel graph types, it is worthwhile to investigate the effects of novelty and familiarity on the attractiveness of information visualizations.

3 *What makes information visualizations usable?*

Research into information visualizations largely concentrated thus far to criteria affecting their usability in terms of accuracy and efficiency (e.g.

Kozhevnikov, Motes, Rasch, & Blajenkova, 2006; Hegarty, Montello, Richardson, Ishikawa, & Lovelace, 2006). Regarding the usability of graphs, studies have mainly used the familiar bar and pie graphs as test materials, in tasks that do not reflect the way information visualizations in mass media are used (e.g. Cleveland & McGill, 1984; Simkin & Hastie, 1987; Spence & Lewandowsky, 1991). This raises the question what features may be responsible for the effectiveness of traditional and more novel information visualization designs as they are used by non-expert users in everyday tasks. In particular, we will study two types of variables.

First, we focus on the perceptual features used to encode quantities. In the literature many different features have been mentioned as being most crucial in reading particular types of graphs, such as length or position along a common scale in reading bar graphs, or angle and area in reading pie graphs. In many empirical studies claims are made regarding the effect of these features on the usability of pie and bar graphs, but these claims are based on researchers' assumptions about which features encode quantity in these graphs, assumptions which not always converge (e.g. Cleveland & McGill, 1984; Simkin & Hastie, 1987; Spence & Lewandowsky, 1991). In this thesis, the role of perceptual features is investigated as perceived by non-expert users. The graphs under study include the familiar bar and pie graphs, as well as many more and less novel types, reflecting the growing variation of popular information visualizations. The effects of these perceptual features on the graphs' usability is tested in tasks reflecting everyday use, in particular comparing the relative magnitude of segments in a graph.

Second, we investigate the possible effect of familiarity on the usability of information visualizations. Information visualizations show a growing variety of designs, which makes the question relevant if these novel types are as usable as the familiar ones.

4 *How do designers and laypeople differ in their understanding and aesthetic preferences?*

Studies into aesthetic experiences of art works have shown differences between experts (with art training) and novices (laypeople) in aesthetic preferences, with novices preferring simple and prototypical stimuli and experts preferring complex and novel stimuli (e.g. McWhinnie, 1968; Reber et al., 2004). Although these studies focus specifically on the appreciation of works of art, in our thesis, we also expect interesting differences between designers and laypeople in the way they understand and appreciate visualizations. Designers are well trained in processing visual information, and may consider familiar designs less appealing because of a lack of originality and visual challenge.

Differences may also be expected based on differences in visual intelligence. Studies have shown that designers have a different kind of visual intelligence than laypeople in art and design (e.g. Blazhenkova & Kozhevnikov, 2010). Designers are supposed to be 'object visualizers' and to pay much

attention to pictorial details and to generate detailed pictorial images of objects and scenes. Spatial visualizers (e.g. engineers) are good at generating schematic images of spatial relations among objects and at imagining spatial transformations. We expect that these differences may result in differences in the way designers and laypeople process and produce information visualizations, and in differences in aesthetic preferences.

1.4 Methodology

In the five studies in this thesis, different methods are used, often in combination. We collect opinions of designers by interviewing them, combined with conducting a literature review (chapter 2). Participants are asked to evaluate the attractiveness of visualizations (chapters 3, 4), to evaluate visualizations' familiarity and perceived ease of use (chapter 4), and to judge the importance of perceptual features of a series of graphs in an online survey (chapter 5) by using Likert and slider scales. They are asked to rank visualizations combined with verbal explanations of their motives for the rankings (chapter 3). Participants are asked to produce information visualizations (chapter 3, 6), and to verbally describe visualizations (chapter 6). Furthermore, they are asked to perform information retrieval tasks with information visualizations in three studies, logging objective performance measures (chapters 3, 4, 5).

We have a few reasons to give priority in this thesis to evaluative methods. The most important is that these methods fit the focus in this thesis on attractiveness, a variable that is best captured by asking participants' judgments directly. It is hard to find valid and feasible objective methods that can capture one's attractiveness judgment better than using fairly simple evaluative measures, i.e., by collecting their behavioral responses in ratings, rankings, and descriptions of likes and dislikes (Palmer, Schloss, & Sammartino, 2013). Apart from ratings, we also collect explicit evaluations by carrying out interviews with designers and collect statements from design handbooks, enabling us to collect their opinions about quality criteria they use (chapter 2). In chapters 4 and 5 participants' perceptions of familiarity, attractiveness and ease of use are measured by asking participants' judgments directly as well. Familiarity is frequently measured by using evaluative methods, for example in psychology (e.g. Blasko & Connine, 1993).

We exploit the full potential of these evaluative methods and at the same time take care to avoid their disadvantages in several ways. In chapter 2 we collect statements from interviews and literature. We avoid subjective interpretations by conducting a survey in which independent participants give their interpretations of the statements, thus providing intersubjective validation of the results. In chapter 3, we combined ratings of aesthetic preferences with a ranking task in which participants were asked to disclose the reasons behind their ranking of graph types. That way, quantitative evaluation data are combined with explicit qualifications given by participants. In chapters 3, 4, and 5

surveys are used. To make sure these results are reliable, a large number of participants was reached by using online surveys distributed via CrowdFlower, a crowdsourcing service of which studies have shown that it yields reliable results (Buhrmester, Kwang, & Gosling, 2011). Lastly, it is a well-known disadvantage of evaluative methods that participants tend to give socially desirable responses. The aesthetic judgments elicited in this thesis can hardly be affected by such bias: judgments were given anonymously and individually; and participants do not need to engage in complicated attitudinal processes connected to controversial issues, they just have to give their evaluation of the attractiveness, familiarity, usability of information visualization designs.

Obviously, the kinds of evaluative methods as we use them have their limitations. Evaluative self-reporting measures do not directly reflect unconscious processes going on when people view and process stimuli. This means that the results of the studies reported here do not allow us to draw conclusions about processing. Other methods are more suitable for that, such as eye tracking (e.g. Goldberg & Helfman, 2011), measuring skin reactions associated with pleasant feelings (Fabrikant et al., 2012) or using fMRI (Aharon, 2001). But our aim is not to measure unconscious processes, but to capture participants' conscious aesthetic judgments. Evaluative methods are most suitable for this and are frequently used in other studies with similar aims and scopes.

In assessing the usability of information visualizations, we combine the evaluative judgment of perceived usability with two standard usability measures (e.g. Cleveland & McGill, 1984; Heer, Kong, & Agrawala, 2009): correctness of performance in information tasks (accuracy) and response time needed to carry out these tasks (efficiency).

Finally, we also use the method of production to gain more insight in the way people use and evaluate visualizations and in differences in understanding and preferences between the two target groups (chapter 3 and 6). Asking participants to visualize information has shown to be a reliable way to obtain insights in how people conceive of information (e.g. Tversky, Kugelmass, & Winter (1991).

Taken together, given the research questions, we consider the methods used in the thesis a suitable way to elicit data from different target groups.

1.5 Overview of the studies

In this section we briefly explain the goal and design of the five studies in this thesis. Three of the five chapters in this thesis (chapters 3, 4, 6) have been published in peer-reviewed journals, two studies (chapters 2, 5) are currently under review. The studies are all self-contained texts, with their own abstract, introduction and discussion section. This results in a certain amount of redundancy in the introductions of the chapters.

Terminology

In the studies we use several terms to refer to the people who form the audience of graphs as they appear in mass media: laypeople, laymen, non-professionals, and the broad or general audience. These differences in terms do not signal differences in target audiences. These differences are the result of differences in referenced literature, and sometimes requests from journal reviewers. The terms refer to the group of people who are not specialists in design, either by education or professional experience. Only in chapter 5, with the term non-experts we refer to the broad audience of people who are not specialized in science or statistics.

Similarly, we refer to the object of study sometimes with the term data visualizations or information visualizations for a broad or general audience, or popular data visualizations. In all cases we refer to familiar and novel types of graphs as they appear in mass media nowadays.

The studies

Chapter 2 is aimed at identifying designers' criteria for good information visualization for a general audience. How important do they consider clarity and attractiveness? Do they intend to communicate objective information or subjective meaning? And do they have ideas about what makes an information visualization attractive? These questions are answered by conducting interviews with professional designers, and by reviewing design literature that is recommended and frequently consulted by designers.

In Chapter 3 we investigate to what extent graphic designers and their audience of laypeople in design share ideas about the clarity and attractiveness of information visualizations. Designers and laypeople are asked to evaluate the clarity, attractiveness, and overall quality of a selection of information visualizations – produced by graphic designers – and to use them in an information retrieval task. Further, they are asked to rank the best and worst 5 graphs, and to explain their motives for the rankings.

In Chapter 4, the influence of familiarity on (perceived) ease of use and attractiveness of information visualizations is investigated. First, we asked participants to assess the perceived attractiveness, familiarity and ease of use of a series of graphs. Second, we asked the same from another group of participants, but then after they had to use these graphs in an information retrieval task.

In Chapter 5 we investigate how perceptual features affect the usability of a series of more and less novel information visualizations. In an online survey, we established which perceptual features are perceived by non-expert users to be most crucial in comparing magnitudes of segments in graphs. In a subsequent study, we asked participants to carry out comparison tasks with the graphs, to assess the effect of perceptual features on the usability of graphs.

In Chapter 6 we study how differences in visual abilities are reflected in performance in visual communicative tasks. For this purpose navigation communication tasks are chosen, since visual-spatial ability has been

associated with performance in navigation communication in several studies. Participants are asked to carry out description tasks based on route images and drawing tasks based on route descriptions. Our aim is to find differential traces of spatial and object abilities in the way designers (object visualizers) and engineers (spatial visualizers) produce and understand visual navigation information.

Chapter 7 contains the general conclusion and discussion.

Information visualization for a general audience: the designer's perspective

Abstract

Graphic designers are increasingly involved in creating 'popular' data visualizations in mass media. We investigated graphic designers' criteria for 'good' design for a general audience by interviewing professional designers and by reviewing information design handbooks. Two issues were central in the study: what is the relative importance of clarity and attractiveness, and what position do designers take in objective representation vs. subjective interpretation of data? Additionally, we investigated what might make an information visualization attractive in the view of designers. According to designers and design handbooks, clarity and attractiveness are the main criteria, with clarity having most emphasis. The designers' aim is to objectively inform the public, rather than communicating personal opinions or concerns. Further, although aesthetics is considered important, design literature hardly addresses characteristics of aesthetics, and interviewees find it hard to define what makes a visualization attractive.

This chapter is based on:

Quispel, A., Maes, A., and Schilperoord, J. (submitted). Information visualization for a general audience: the designer's perspective.

2.1 Designing information visualizations for a general audience

This explorative study takes the perspective of the producer of mass media information visualizations and addresses three issues that hitherto have only been amply discussed in theoretical and applied literature on information visualization¹. What is the role and relative importance of clarity and attractiveness? Do designers aim to present objective information or convey subjective meaning? And what makes an information visualization attractive? The study is based on a collection of statements and opinions with respect to these issues which have been derived from two sources: interviews with professional designers and information design handbooks which are recommended and frequently consulted by information designers.

Traditionally, information visualization techniques have been developed mainly for science and statistics, and were first and foremost meant to allow expert users to explore and analyze data quickly and accurately. In the past decades, however, many collections of data have become freely available, and so has software for data visualization. As a result, an increasing number of designers has started to apply visualization techniques to create data visualizations for popular purposes. At the same time, the intended audiences of those visualizations has expanded from expert users to include various groups of lay users as well (Vande Moere & Purchase, 2011). The increasing popularity of data visualizations is also testified by the growing number of books for non-expert users that provide guidelines for data visualization to be used by non-scientific readers and showcase a huge variety of popular data visualization techniques (e.g. Klanten, 2008; McCandless, 2009). Among the designers involved in creating popular data visualizations are many graphic designers and other types of designers such as interaction designers, who have been educated at art and design academies (e.g. <http://www.catalogtree.net>; <http://lust.nl>; <http://tulpinteractive.com>). Novel ways of visualizing data have been developed for business, government, newspapers, magazines, and internet platforms. These popular forms of data visualization are not only meant to allow an efficient and accurate reading of the data, as in science, but also to inform broad audiences about facts and developments in society. Consequently, designing such data visualizations may call for other design criteria than the ones applied to design graphs meant to serve scientific, analytical purposes. In this study, we are especially interested in those other criteria.

Some researchers in the scientific field of information visualization assume that aesthetic criteria are a key factor in communicating quantitative

1 In the scientific field of information visualization, the term information visualization refers to the computer-aided interactive visualization of big data with the aim to amplify cognition (Card, MacKinlay, & Shneiderman. 1999). In graphic design, the term refers to a broader category of visualizations meant to inform or instruct people, including data visualization. In this article with the term information visualization we refer to the visualization of quantitative data, be it limited or big, static or interactive.

information to broad audiences (e.g. Judelman, 2004; Kosara, 2007; Lau & Vande Moere, 2007; Vande Moere & Purchase, 2011). Those researchers propose collaborations between scientists and designers to strike an adequate balance between information value on the one hand and aesthetics on the other, i.e. between *clarity* and *attractiveness*. They thus implicitly assume that aesthetics is part of the expertise of designers. Some also seem to suggest that designers of popular data visualizations tend to put more emphasis on attractiveness than on clarity. As Vande Moere and Purchase put it: '(...) complex and socially relevant issues might best be communicated to a large audience through popular media using an artistic and engaging visualization (even if its designer knows that such a method is not the most effective or efficient).' (p.361). According to Kosara (2007) 'artistic' and 'pragmatic' forms of visualizations even seem irreconcilable: 'Visual efficiency does not play a role in artistic visualization, quite the contrary. The goal is not to enable the user to read the data, but to understand the basic concern.' (p.634). Related to this latter quote, is the assumption made by several researchers that designers intend to convey subjective meaning underlying the data, rather than objectively presenting data and facilitating insight in the data (e.g. Kosara, 2007; Lau & Vande Moere, 2007). Designers would employ 'ambiguous and interpretative methods' in order to engage the user and provoke personal reflection (Gaver, Beaver, & Benford, 2003), and their designs are supposed to involve subjective decisions and stylistic influences, and to be highly interpretative (Lau & Vande Moere, 2007).

The assumptions as described above reflect the way some scientists in the field of information visualization and human-computer interaction think about characteristics of popular information visualizations. But how do designers themselves think about these matters? In the present study, we take the perspective of the designers as the producers of popular data visualizations as a starting point. To shed light on these matters, we interviewed ten professional designers. In addition, several handbooks on data visualization were reviewed in search for criteria for data visualization for a general audience. The selected handbooks can be assumed to reflect 'best practices' in information design because they are recommended by the International Institute for Information Design (IIID); an authoritative institution in the field of information design; and because they are frequently consulted by designers.

In particular, we focused on the following questions:

How do designers look upon the relative importance of clarity and attractiveness in information design?

Any message, be it casted verbally or visually, is created with the objective to be understood by the intended audience. Therefore, the message should first and foremost be clear and understandable. At the same time, however, in order to be notified in the first place, the message must attract and hold attention; in other words the message has to be attractive for the intended audience. In this study, we consulted designers about what they consider to be the most

important criterion in designing visualizations. Do they consider clarity to be more important or attractiveness? Or do they consider clarity and attractiveness equally important?

What position do designers take in objective representation of data vs. providing subjective interpretation in visualizing information?

When people use language, they have many conventional signals at their disposal to differentiate between expressing objective vs. subjective content (modal verbs, different types of connectives, etc.). When designers communicate messages using the visual modality, they arguably also aim at expressing either objective knowledge or subjective interpretations of it. However, the visual modality may not have a conventionalized set of signals to mark the difference between facts and opinions. So, the question is whether designers differentiate between these two types of information in their designs, and, if so, how they mark this distinction. With the interviews and literature review, we wanted to investigate the opinions of designers with respect to this objective-subjective issue.

What do designers consider the defining characteristics of attractiveness in information design?

Scientific studies in information visualization usually do not offer explicit hypotheses about what might make an information visualization attractive. Aesthetics is a very complex notion, and several theories and studies have been constructed and conducted about features of aesthetics (e.g. Hekkert, 2006; Reber, Schwarz, & Winkielman, 2004). Some studies start from the theoretical assumption that *novelty* (or related notions like *originality*, *innovativeness*, or *uniqueness*) is a factor causing attractiveness (e.g. Lavie & Tractinsky, 2004), while others state that especially experts (in art) are attracted to novel and complex, instead of familiar and simple stimuli (McWhinnie, 1968; Bourdieu, 1987; Gombrich, 1995). Other theories suggest that attractiveness is a matter of striking a balance between, for example, novelty and familiarity, or between simplicity and complexity (e.g. Berlyne, 1971). And yet other studies conjecture that attractiveness results from familiarity (resulting from repeated exposure) and experienced ease of use (Zajonc, 1968/1984; Reber et al., 2004), or from specific design features such as being embellished or abstract (e.g. Levy, Zacks, Tversky, & Schiano, 1996) or using certain color palettes (Fabrikant, Christophe, Papastefanou, et al., 2012). We wondered what the designers' views would be on what characteristics contribute to attractiveness of information visualizations.

The first two questions represent scales with dichotomous terms discussed by researchers as being important criteria for 'good design': clarity vs. attractiveness, and objectivity vs. subjectivity. In the remainder of the chapter we will refer to these scales using the first term (clarity, objectivity). The third question is an open question concerning features of attractiveness in information visualization.

We collected answers from designers in two ways. First, we interviewed 10 professional designers who are regularly involved in information visualization for broad audiences about their standards and ways of working. Second, we selected relevant fragments from recommended and frequently consulted design handbooks.

2.2 Interviews and literature review

2.2.1 Method

Interviews

10 semi-structured interviews were conducted with professional designers.

Participants. Interviewees were 10 professional designers who are used to design data visualizations for general audiences. Their educational backgrounds were: graphic design (4), interaction design, computer sciences, industrial design, journalism, mathematics & graphic design, and journalism & industrial design. They were selected for at least one of the following three reasons: (i) having been rewarded with prestigious design prizes (Infographics Jaarprijs (7), Dutch Design Award (1), Malofiej Award (1)); (ii) leading the infographics department of a Dutch national newspaper (2); and (iii) being regular speakers at information design conferences (7), such as the Infographics Jaarcongres (Dutch yearly infographics conference).

Procedure. The designers were approached first by e-mail and then by telephone. They were informed that the goal of the interview was to gain more insight in their working methods and criteria for good design; that the interviews were part of an ongoing PhD research, and that the results would be published anonymously. Interviews took about 45 to 60 minutes, and were conducted in face to face settings. Six designers were interviewed individually while two interviews were held with two designers who work closely together. The interviews were recorded and later transcribed.

Each interview consisted of three parts. The first part was meant to get acquainted with the interviewees and their design practice. Interviewees were asked about their ways of working when designing information visualizations for a general audience. In the second part they were asked the main questions regarding clarity, objectivity, and attractiveness. Questions were framed in a general way to grant interviewees the opportunity to mention other elements than anticipated. In the third part, interviewees were asked if they ever conducted usability tests, if they checked data, what they thought about misleading graphs, and what design literature they consulted. Finally some typical examples from the designer's portfolio were examined to further illustrate their ideas. The full list of questions appears in Appendix A.

Handbooks

26 information design handbooks were reviewed. The full list of references is in Appendix B.

Selection. The selection of handbooks was based on at least one of the following three criteria. First, general information design handbooks were selected which according to the International Institute for Information Design (www.iiid.net, June, 2015) can be considered 'important basic resources', a qualification which is 'based on evaluations of experts from America, Asia and Europe'. From this source we selected all handbooks which focus on information design (n=8) (and not for example on architecture or urban design). Second, we selected all design handbooks that the interviewees had reported to consult regularly (n=12). And third, we selected the 15 information design and data visualization books that, according to library loan statistics (January 2006 – June 2015), are most frequently consulted by design students at the AKV|St. Joost academy of art and design, Avans University (the first author's affiliation). In sum, this search resulted in 26 design books that were reviewed (see Appendix C for full references).

Review procedure. The indexes and tables of content were inspected for relevant terms (aesthetics, attractiveness, clarity, principles, and related criteria terms). If they were found, the referenced sections were reviewed. Because explicit references to the terms as mentioned were hardly ever found in the indexes and tables of content, we also consulted the introductory sections and chapters of the handbooks. In those cases where a handbook was divided into parts that contained separate introductions, we reviewed the introductions of each part.

2.2.2

Analyses

Transcriptions and selected excerpts from handbooks were analyzed in three steps. First, we selected all statements which somehow appeared to include an answer to one of the three questions or expressed an opinion with regard to these questions. This way the materials were made manageable in terms of size and focus. Next, we presented statements concerning the first two questions, i.e. clarity and objectivity, in a forced choice task to respondents, in order to collect intersubjective validation of the designers' position on these two questions. Finally, we conducted an additional qualitative analysis of the fragments, including those addressing the third question concerning attractiveness, thereby taking into account the data collected in the previous steps.

Step 1: selecting relevant statements

Interviews. The interviews yielded 25 statements and excerpts that explicitly addressed the itemized issues: 9 on the importance of clarity and attractiveness, 9 about objectivity vs. subjectivity, and 7 about features contributing to attractiveness.

Handbooks. From the 26 selected handbooks statements were taken that contained explicitly marked normative expressions with regard to the criteria and objectives of information visualization design. These statements typically contain modal verbs such as 'must' or 'should' in relation to information design, e.g. 'Information displays should be [...]', or normative markers such as 'the first priority of an information designer is [...]'; 'principles of analytic design: [...]'; 'the purpose of visualization is [...]', or 'excellence consists of [...]'. This selection resulted in 28 normative statements coming from 22 handbooks (4 handbooks contained no normative statements regarding criteria and objectives). Of these statements 19 addressed the clarity question, 7 the objectivity question, and 2 the attractiveness question.

Step 2: explorative survey

Before the statements from the interviews and design handbooks were analyzed, a separate survey was conducted in order to obtain an intersubjective agreement on the interpretation of the 44 clarity and objectivity statements. The statements were presented to an independent group of participants in a forced choice task survey. The survey was constructed in Qualtrics and distributed via email to teachers of various disciplines and master of design students of Avans University. The survey was taken this way by 39 respondents (22 designers, 17 design lays), all with higher vocational or university level education. For each statement, these independent participants were asked to judge the position of the statement's author on a dichotomous scale (clarity or attractiveness, subjectivity or objectivity. Figure 1 shows an example of a clarity statement as it was presented in the survey.

Which criterion is the most important according to the quote:
clarity or **attractiveness**?

"You hope it is clear, and that people will use it."

clarity	attractiveness
---------	----------------

Figure 1 Example of statement and scale from the interpretation survey

Table 1 shows the results from the survey (See Appendix B for all statements and scores per statement.) A statement was considered to represent a certain position on the scale (e.g., either clarity or attractiveness), when at least two-third of the respondents selected this position (66.66% or more). Otherwise, the statement was considered undecided.

Nr of statements	Nr of statements in which each criterion is judged as most important		
	Clarity	Attractiveness	Undecided
Interviews: N = 9	6	3	0
Handbooks: N = 19	14	0	5
Nr of statements	Objective	Subjective	Undecided
	Interviews: N = 9	2	2
Handbooks: N = 7	4	0	3

Table 1 Number of statements in which each criterion is judged as most important

The results show that in the majority of the 'clarity' statements from the interviews and the handbooks clarity is considered to be the most important criterion (20/28). Moreover, there is broad agreement on this choice, as is shown in Appendix C, listing the percentages of the respondents choosing clarity as main criterion. In 16 of the 20 cases more than 80% of the respondents agreed on this interpretation of clarity being the main criterion.

As for the objectivity – subjectivity statements, results are undecided for half of the statements (8/16). In 6 of the remaining 8 cases the statements are interpreted as objectivity being the main communicative goal. Only a small minority of statements (2/16) is interpreted as subjectivity being the main goal.

Qualitative analysis

In this section we answer the three research questions by analyzing the answers and fragments addressing them. To exemplify analyses, several statements will be provided with relevant passages put in italics and each of them accompanied by the source (I=interviews; H=handbooks) and the percentage of respondents in the survey (as given in Appendix C) which agreed on its interpretation.

Clarity vs. attractiveness

With respect to the main criteria for good information visualizations for a general audience, handbooks and interviews offer a consistent picture: clarity and attractiveness are considered by far the most important criteria. Of these two, clarity is considered the most important criterion, both in the interviews and in the handbooks. In interview statements, it is often mentioned first; given more emphasis and often more elaborated upon, as is testified by (1).

- 1 'Relevance. [...] Meaning the infographic has to contribute to conveying information. *It must be clear.* That means you have to make choices with which you guide your reader through the infographic. Accessibility. The reader must not give up because of complexity or because he doesn't know where to begin. *And attractiveness.* [...] It must surprise, excite, make curious.' (I; clarity: 96.4%)

Also, when attractiveness is mentioned as being important, statements to this effect are immediately moderated by a but- or however-phrase stating that clarity is more important. See the following statements:

- 2 'I think you must be able to see what the subject is. Quickly see where you have to search. Not simplify by leaving out, but clarify by layering. *Attractiveness tops. But I am not an artist.* Form follows content. If the image is pretty but non-informative, than that is not sufficient for me. Not more image than content.' (I; clarity: 100%)
- 3 'The purpose of visualization is insight, not pictures. A visualization's function is to facilitate understanding. Form has to follow this function. This does not mean that aesthetics are not important – *they are.* [...] *However,* it is not only aesthetics that help to increase the information flow.' (H: Scheiderman, in Klanten, 2010, p.8; clarity: 82.1%)

In other cases, the priority is obvious, since the statements only mention efficiency or clarity (e.g. 4):

- 4 'Information design as a discipline has the efficient communication of information as its primary task.' (H: Wildbur & Burke, 1998, p.6; clarity: 96.4%)

In only 3 of the statements (all from the interviews) attractiveness is considered most important, as in (5).

- 5 'Clarity and aesthetics. Where the emphasis is put, depends on the target group and the assignment. If it is for a newspaper for a broad non-expert audience, you have to tell a story that is engaging; then it is not all about efficiency.' (I; clarity: 21.4%)

In some statements, attractiveness and clarity are mentioned as equally important, which results in the interpretations being undecided, as in the following cases:

- 6 'The only conclusion possible is that design always involves three inextricably related elements, however much their relative proportions may differ from one application to the next, namely: durability, usefulness, and beauty.' (H: Mijksenaar, 1997, p.18; clarity: 46.4%)
- 7 'The optimum synthesis of aesthetics and information value remains the essential objective in every type of diagrammatic presentation.' (H: Herdeg, 1981, p.6; clarity: 64.4%)

Several interviewees and design handbooks also stress that striving for clarity does not mean that information should be simplified by leaving things out. Instead, complex information should be made accessible and understandable by design means such as layering: a visual ordering of information that allows both overview and detailed reading, and that highlights what is important (e.g. Few, 2004; Tufte, 1990).

Despite the value they assign to clarity, most of the interviewees indicate in their answers to the additional question in this regard, that they hardly ever test if their audience understands their designs. If they test their designs before publication at all, this is usually done informally among fellow designers.

Objectivity vs. subjectivity

The second main question was if designers aim to communicate objective or subjective information. The dichotomy objective vs. subjective suggests a sharp contrast between the two, but the answers and fragments addressing this issue show a more nuanced picture. In the statements subjectivity does not mean that readers are forced to swallow the designer's truth or opinions. Rather, it covers the idea that most of the designers aim to do more than just present data. They feel they need to add elements enabling viewers to arrive at an adequate or intended interpretation of the data. More than in the previous questions, the interpretation of the statements in this section is undecided (8/16). See for example the following statements:

- 8 'Complete objectivity doesn't exist, you always make choices. But it is not our aim to give our personal opinion.' (I; objectivity: 60.7%)
- 9 'I do not give my own judgment. You have to interpret. But people have to make their own judgment.' (I; objectivity: 57.1%)
- 10 'You have to interpret; simply provide data is pointless. But people have to make their own judgment.' (I; objectivity: 28.6%)

Complete objectivity would mean presenting raw data, which is something designers feel is of no use for non-expert users. Especially for these audiences the designer needs to choose which data are relevant and which patterns and relationships to show in order to convey the intended message. Furthermore, the designer needs to choose design elements that help to explain what the visualization shows, such as labels, pointers, line width and colors that direct the eyes to what is important (Yau, 2011). 'Making choices' in (8) and 'interpret' in (9) and (10) therefore most likely refer to this kind of decisions and choices. Apparently, survey respondents had difficulty to decide if these statements should be interpreted as intending to be objective (no opinions, no judgments) or subjective (interpret). While these statements seem to have similar meaning ('interpretation, but no judgment'), the interpretation of (8) and (9) is rather

undecided (60.7% and 57.1% objectivity respectively), whereas (10) is interpreted as subjective (28.6% objectivity). It may be the case that the difference between undecided and subjective results from the way the word 'interpret' is embedded in the sentences. In (8) and (9) subjectivity ('you always make choices' and 'give my own judgment') is construed as inevitable but explicitly denied to be the aim of designs, or it is first denied, followed by the need to interpret (9; 'subjectivity'). In statement (10) on the other hand the need to interpret is mentioned first. Besides these undecided cases, 6 of the 16 statements were interpreted as objectivity, as in (11):

- 11 'Show the data; [...] avoid distorting what the data have to say.' (H: Tufte, 2001, p.13; objectivity: 92.9%)

Only 2 of the 16 statements were interpreted as subjectivity being the main goal. An example is (12) in which the terms 'engage', 'feeling', and 'atmosphere' are mentioned and apparently associated with subjectivity.

- 12 'Then you must be able to experience the story. Make the story manifest. Engage people. [The question is:] how to translate a feeling, an atmosphere, into something visual?' (I; objectivity: 10.7%)

Apart from the question what subjectivity means exactly (ranging from allowing readers their own interpretation of the data to conveying personal opinions on the part of the designer), an interesting other question is whose interpretation is expressed in the designs. For example in (13), it is suggested that it is not the designer's personal opinion that counts, but his clients' message:

- 13 'I try to make things visual as soon as possible, and then discuss with the client what exactly it is they want to tell. It is not about my own message.' (I; objectivity: 64.3%)

This response raises two issues: the source and correctness of the data, and the relationship between what the data have to say and the message that the client wishes to convey. From the answers to the additional questions in the interviews, it becomes apparent that designers are well aware of these issues, and of their responsibility in the way information is interpreted and visualized. As for the source and correctness of the data, in many cases the data are supplied by the client (business, government, or editorial), but in 4 of the interviewees' practices it regularly occurs (mostly in journalism) that the designers collect the data themselves. In all those cases, designers emphasize that they assign much value to correctness of the data. If they have access to the sources, they check them, and if they come across mistakes, omissions or inconsistencies, they contact the client and correct them. When it comes to the interpretation of the data, there are roughly two scenarios. In 6 of the interviewees' design practices the situation occurs that the client has no clear

message to convey. Especially governmental and business clients sometimes supply data, but leave it to the designer to interpret them. In these cases the designer discusses with the client what s/he believes to be the most important information and what s/he thinks the message is that the client might want to convey. See for example the following statement:

- 14 'Sometimes the message is clear, but not always. These days I am increasingly employed by businesses, and then the assignment is often very vague. I take an active role then to find out: what is the message? They often don't know. (...) Or governments have made analyses and want to show something about them to managers or to the public. Then they come to us with piles of reports full of important information: can you make a graphic of this? (...) Then we have to extract the essence ourselves.'

In the other scenario, also mentioned as occurring regularly in half of the interviewees' design practices, the client has a clear conception of the message s/he wants to transmit. In these cases the designers all state that this message has to be in accordance with the data provided. None of them are prepared to 'lie' or deceive with information visualizations by manipulating data, distorting scales or whatever means, and clients, they say, hardly ever ask them to purposely do so. See for example statement 15:

- 15 'It only happened once, with a big international non-profit organization (...). When the data did not fit the story, they would just make another selection from the data. I will never work for them again.'

In another example the designer was asked not so much to 'lie', but to upscale an organization's role in a decision process as being central, while in reality, to go by the data, their role was only marginal. The designer's reaction:

- 16 'When something is not right in a graphic, people notice that quicker than in a text. (...) You cannot visualize what is not there. [Referring to the example:] We reached a compromise: the organization on top instead of at the bottom, the process reversed. If I had given them what they wanted and had put them literally in the center, it would have become a very bad visualization with a weird twist in it, and people notice that.'

In this designer's opinion, it is hard to deceive viewers with visualizations. Many researchers and designers would disagree with this, such as Tufte (1983), who showed excellent examples of how to 'lie' with graphs in ways that are still ubiquitous. We will not further elaborate on this matter in the framework of this study, but it is an interesting research question to what extent people are visually literate enough nowadays to not let themselves be tricked by distorted or manipulated data visualizations.

In the case of newspapers-as-clients, designers sometimes find themselves in a position where editors reject a graphic if it does not fit the story they

intend to tell. Also then, the designers report they refuse to create a graphic that does not fit the data.

Attractiveness features

We also searched in the handbooks and interviews for fragments that offer information on what determines or characterizes attractiveness. The results illustrate how difficult it is to put into words what makes a design – or any artifact – attractive. See for example the following statements from the interviews:

- 17 'That is in the design. I think in metaphors, which I try to use as an illustrative element. [...] I use associations between subject and form [...]. It is mainly about information density.'
- 18 'That follows from the process. It designs itself. Beauty that you see in it is a bit of intuition that you are doing alright. I don't think we have a visual language of our own.'
- 19 'We start from what we consider good ourselves. That is hard to define. Data visualization is usually clear in terms of archetype. Then look at contrast, define archetype, and then the subject, and then you come in an atmosphere, and aesthetics. Aesthetics is important. We do what we like.'

None of the interviewees suggests that attractiveness might be found in familiarity, as is assumed in several theories (Zajonc, 1968/1984; Reber et al., 2004). Some, however, refer to notions such as novelty (something unique, or surprising), as in the following examples:

- 20 'In interactive visualizations: playful movement, something surprising... Use of color... More feeling. Just as much information, but beautifully designed.'
- 21 '(...) something that is unique and tells a story, that attracts attention and is remembered.'

In the design handbooks, explicit statements concerning features that might define attractiveness turn out to be hard to find, despite the importance that is attached to attractiveness. Two statements seem to attempt to define what makes an information visualization attractive:

- 22 'Graphical elegance is often found in simplicity of design and complexity of data.' (Tufte, 2001, p.177)
- 23 'Elegance is a measure of the grace and simplicity of the designed product relative to the complexity of its functions.' (Herdeg, 1981, p.8)

These statements are in accordance with theories describing attractiveness in terms of a balance between extremes, in this case between simplicity and complexity. Furthermore, this idea that people would be attracted to 'simplicity in complexity' resembles an assumption in simplicity theory stating that people find it pleasing when seemingly difficult information is surprisingly easy to understand (Chater, 1999). This suggests that attractive information design results not from simplifying things, but from clarifying complex information, as is also reflected in one statement from the interviews concerning the importance of clarity:

24 '(...) Not simplify by leaving out, but clarify by layering. (...)'

and in one interviewee's response to the question concerning features of attractiveness:

25 '(...) many data that show patterns. (...)'

In sum, novelty and simplicity in complexity seem to play a role in attractiveness, but taken together, the statements from the interviews and the design handbooks offer too little explicit information to draw conclusions about what designers consider features determining attractiveness of information visualizations.

2.3 Discussion and conclusions

According to designers, data visualizations for non-expert audiences should be attractive and, most importantly, be clear. Contrary to what is sometimes conjectured by scientists, designers do not put more emphasis on appearances than on understandability. On the contrary, attractiveness is considered important, but clarity is paramount. Interestingly, despite the importance they assign to clarity, the designers indicate in the interviews that they hardly ever test their designs among the intended audiences. If they test them at all, they usually do this among fellow designers. Therefore, it would be interesting to test if popular data visualizations are indeed understood by a general audience of non-designers.

Further, it becomes apparent from the design handbooks and especially the interviews that designers act very responsibly in the way they visualize information. They attach much importance to correctness of data, and they are careful not to deceive their audience. Messages conveyed through visualizations should fit the data. They feel no need to convey personal opinions or judgment, but they do feel that they need to help their readers to interpret visualizations. Still, it would be interesting to know if it is at all possible to convey personal opinions in an information visualization. Some scholars, like Yau (2011) and Vande Moere and Purchase (2011) argue that popular forms

of data visualizations are employed not to present cold facts, but to convey concern, or emotion, referring to popular visualizations such as *We feel fine* (Kamvar & Harris, 2009). Then it is the question whether this is mainly a matter of content choice (e.g. emotional terms used on Twitter or Facebook, or environmental issues), or if this is (also) achieved by design means. And if the latter is the case, how exactly do designers apply what kind of design means to express emotion or personal beliefs in an information visualization? This would be an interesting direction for future research.

Concerning attractiveness, it is striking how little information can be found about what might contribute to the attractiveness of visualizations. Hardly any information can be found on this matter in design handbooks, and designers find it hard to describe what features might make their designs attractive. This is not surprising, of course. Also in other disciplines, such as literature, it will be hard to find discourse that explains what makes a text or some other artifact attractive, and it will be equally hard for other practitioners to put into words what makes their works appealing. Designers mention features such as novelty, which might be expected, considering that it is a designer's job to create new things. The scarce handbook fragments point in an interesting direction concerning simplicity in complexity. But in all, the fragments and statements are too few in number to base conclusions on. Yet, designers and both scientific and design literature agree that aesthetics play an important role in information visualization, equally important to understandability, and research has also shown interactions between aesthetics and (perceived) usability. It would therefore also be an interesting direction for future research to further investigate the characteristics of aesthetics in information visualization, and its relationships with usability.

Acknowledgements

The authors cordially thank the following information designers for their generous cooperation in the interviews: Erik van Gameren, Daniel Gross, Remy Jon-Ming, Joris Maltha, Dimitri Nieuwenhuizen, Frédéric Ruys, Yassine Salihine, Karin Schwandt, Eugene Tjoa, Jan Willem Tulp.

Appendix A

Interview questions

Introductory questions

- Can you briefly describe your educational background and professional career / experience?
- What kind of clients do you usually work for and prefer to work for? (e.g. editorial, government, business)
- Can you describe a typical work process? How are you provided with the data, who is responsible for the analysis, who constructs the message, and how do you communicate with the client about the design?

Main questions

- What are the main criteria for a good data visualization for a general audience?
- Do you show opinions in your data visualization designs?
- What makes a data visualization attractive?

Additional questions

- What is your opinion about misleading with graphs? Do clients ever ask you to lie with graphs, and have you ever done that? Do you check the data you are supplied with?
- Do you test your designs before they get published?
- Do you consult design handbooks, and if so, which?

Appendix B

Reviewed design handbooks:

LL = library loan

IW = interviewees' literature

IIID = recommended by IIID

- Bertin, J. (2011). *Semiology of Graphics: diagrams, networks, maps*. Redlands, CA: ESRI. [IW]
- Brückner, H. (2004). *Informationen Gestalten*. Bremen: Hauschild. [LL; IIID]
- Cairo, A. (2013). *The functional art: an introduction to information graphics and visualization*. Berkeley, CA: New Riders. [IW]
- Few, S. (2004). *Show me the numbers: designing tables and graphs to enlighten*. Oakland, CA: Analytics Press. [IW]
- Harris, R. (1999). *Information graphics: a comprehensive illustrated reference*. Atlanta, GA: Management Graphics. [IW]
- Herdeg, W. (ed.) (1981). *Diagrams: the graphic visualization of abstract data*. Zürich: Graphis Press. [LL]
- Holmes, N. (2005). *Wordless Diagrams*. New York: Bloomsbury. [LL]
- IIID Japan (ed.) (2005). *Information Design Source Book*. Basel: Birkhauser-Publishers for Architecture. [IIID]
- Jacobson, R.E. (ed.) (1999). *Information design*. Cambridge, MA: MIT Press. [IIID]
- Klanten, R. (ed.) (2008). *Data Flow - Visualising Information in Graphic Design*. Berlin: Gestalten. [LL]
- Klanten, R. (ed.) (2010). *Data Flow 2: visualizing information in graphic design*. Berlin: Gestalten. [LL; IW]
- Klanten, R. (ed.) (2011) *Visual Storytelling: inspiring a new visual language*. Berlin: Gestalten. [LL]
- Lima, M. (2011). *Visual complexity: Mapping patterns of information*. New York: Princeton Architectural Press. [LL]
- Maeda, J. (1999). *Design by numbers*. Cambridge, MA: MIT Press. [IW]
- McCandless, D. (2009). *Information is beautiful*. London: Collins. [LL; IW]
- Mijksenaar, P. (1997). *Visual Function: an introduction to Information Design*. Rotterdam: 010 Publishers. [IIID; IW]
- Neurath, O. (2010). *From hieroglyphics to Isotype: a visual autobiography*. London: Hyphen. [IW]
- Tufte, E. R. (1990). *Envisioning Information*. Cheshire, Connecticut: Graphics Press. [LL; IIID]
- Tufte, E. R. (1997). *Visual Explanations*. Cheshire, Connecticut: Graphics Press. [LL; IIID]
- Tufte, E. R. (2001). *The Visual Display of Quantitative Information*. Cheshire, Connecticut: Graphics Press. [LL; IIID; IW]
- Tufte, E.R. (2006). *Beautiful Evidence*. Cheshire, Connecticut: Graphics Press. [LL]
- Ware, C. (2004). *Information Visualization - Perception for Design*. San Francisco: Morgan Kaufmann. [IW]
- Ware, C. (2008). *Visual Thinking for Design*. Amsterdam: Elsevier/Morgan Kaufmann. [IW]
- Wildbur, P. (1989). *Information Graphics*. Houten: Gaade. [LL]
- Wildbur, P. & Burke, M. (1998). *Information Graphics - Innovative Solutions in Contemporary Design*. London: Thames and Hudson. [LL; IIID]
- Woolman, M. (2002). *Digital Information Graphics*. London: Thames & Hudson. [LL]

Appendix C

Statements and their scores (percentages of choices for clarity and objectivity)

I = statement from interview

H = statement from design handbook

Statements on clarity vs. attractiveness

Statements judged to express clarity as main criterion (> 66.66%)		% clarity
I	'I think you must be able to see what the subject is. Quickly see where you have to search. Not simplify by leaving out, but clarify by layering. Attractiveness tops. But I am not an artist. Form follows content. If the image is pretty but non-informative, than that is not sufficient for me. Not more image than content.'	100.0
H	'To communicate quantitative information effectively first requires an understanding of the numbers, then the ability to display their message for accurate and efficient interpretation by the reader.' (Few, 2004, p.10)	100.0
H	'Our goal is to enable the user to understand and find his way [...]. Accuracy always takes priority over esthetics.' (Brückner, 2004, p.11)	100.0
I	'Relevance. [...] Meaning the infographic has to contribute to conveying information. It must be clear. That means you have to make choices with which you guide the reader through the infographic. Accessibility. The reader must not give up because of complexity or because he doesn't know where to begin. And attractiveness. [...] It must surprise, excite, make curious.'	96.4
I	'It must be legible. And the reader must be able to read his own story: can you compare things, zoom in on details, etc. You often see beautiful images with a lot of data without the story being clear. That is not good.'	96.4
H	'A graphic designer is expected to convey a message as clear as possible by creating order in text and image. Information design is a spectrum of design that is mainly occupied with giving consumers information in the clearest and most direct manner.' (Wildbur, 1989, inside cover)	96.4
H	'The challenge is to develop ways of arranging the most relevant data in the clearest manner and the smallest amount of space.' (Woolman, 2002, p.11)	96.4
H	'Information design as a discipline has the efficient communication of information as its primary task.' (Wildbur & Burke, 1998, p.6)	96.4
H	'Excellence in statistical graphics consists of complex ideas communicated with clarity, precision, and efficiency.' (Tufte, 2001, p.13)	92.9

	Statements judged to express dominance of clarity as main criterion (> 66.66%)	% clarity
H	'Its [information design] purpose is the systematic arrangement and use of communication carriers, channels, and tokens to increase the understanding of those participating in a specific conversation or discourse.' (Jacobson, 1999, p.4)	89.3
I	'You hope it is clear, and that people will use it.'	85.7
H	'The first goal of an infographic is not to be beautiful just for the sake of eye appeal, but, above all, to be understandable first, and beautiful after that, or to be beautiful thanks to its exquisite functionality. A good graphic realizes two basic goals: it presents information, and it allows users to explore that information.' (Cairo, 2013, p.XX)	85.7
H	'Although aesthetics are taking on an increasingly important role, we must always ensure that visualizations make things easier to understand.' (Richli in Klanten, 2008, p.185)	85.7
H	'The goal of information design must be to design displays so that visual queries are processed both rapidly and correctly for every important cognitive task the display is intended to support.' (Ware, 2008, p.14)	85.7
H	'The purpose of visualization is insight, not pictures. A visualization's function is to facilitate understanding. Form has to follow this function. This does not mean that aesthetics are not important – they are. [...] However, it is not only aesthetics that help to increase the information flow. Narrative is a very powerful tool as well.' (Schneidermann in Klanten, 2010, p.8)	82.1
H	'The idea is to make designs that enhance the richness, complexity, resolution, dimensionality, and clarity of the content.' (Tufte, 1997, p.9-10)	82.1
I	'It must tell a story. Provide insight into something. No info no graphic. Aesthetics is also important, if it does not distract.'	75.0
I	'The data set is important. It must tell something, provide a different insight than with Excel. It starts with the analysis of what you could give insight in.'	71.4
H	'Certain [design] choices become compelling because of their greater efficiency.' (Bertin, 2011, p.9)	71.4
	Statements undecided (66.66 – 33.33%)	% clarity
H	'When consistent with the substance and in harmony with the content, information displays should be documentary, comparative, causal and explanatory, quantified, multivariate, exploratory, skeptical.' (Tufte, 1997, p.53)	64.3
H	'The optimum synthesis of aesthetics and information value remains the essential objective in every type of diagrammatic presentation.' (Herdeg, 1981, p.6)	64.3
H	'Too many data presentations, alas, seek to attract and divert attention by means of display apparatus and ornament.' (Tufte, 1990, p.33)	60.7
H	'The only conclusion possible is that design always involves three inextricably related elements, however much their relative proportions may differ from one application to the next, namely: durability, usefulness, and beauty.' (Mijksenaar, 1997, p.18)	46.4
H	'Combining beauty and truth, they [data visualizations] are, at their best, inspiring, fascinating, visually interesting and easy to read, while conveying complex levels of information in an impactful way.' (Losowsky in Klanten, 2011, p.6)	39.3

Statements judged to express dominance of attractiveness as main criterion (< 33.33%)		% clarity
I	'It starts with the data. A design cannot transcend the content. [...] Then you must be able to experience the story. Make the story manifest. Engage people. It must contain something intriguing that attracts.'	25.0
I	'Clarity and aesthetics. Where the emphasis is put, depends on the target group and the assignment. If it is for a newspaper for a broad non-expert audience, you have to tell a story that is engaging; then it is not all about efficiency.'	21.4
I	'Data must be correct, fit the story, and it must provide an extra or different insight into the story that accompanies it. You don't have to be able to see immediately what it is about - that is too difficult with some financial constructions. But I do try to make something stand out, as a sort of hook.'	10.7

Statements on objectivity vs. subjectivity

Statements judged to express dominance of objectivity as main goal (> 66.66%)		% objectivity
H	'Show the data, [...] avoid distorting what the data have to say.' (Tufte, 2001, p.13)	92.9
H	'The first priority of information design is the correct communication of serious subject matter.' (Brückner, 2004, p.7)	89.3

Statements judged to express dominance of objectivity as main goal (> 66.66%)		% objectivity
H	'A good graphic realizes two basic goals: it presents information, and it allows users to explore that information.' (Cairo, 2013, p.73)	85.7
H	'Show comparisons, contrasts, differences. Show causality, mechanism, explanation, systematic structure.' (Tufte, 2006, p.127)	75.0
I	'I think the essence is that you confine yourself to facts and figures and let the reader draw his own conclusions. I don't feel the need to give my own viewpoints, I rather do it the other way around.'	71.4
I	'We don't need to bring our own truth, we rather map everything, so that people can form their own opinion.'	67.9

Statements undecided (66.66 – 33.33%)		% objectivity
I	'I try to make things visual as soon as possible, and then discuss with the client what exactly it is they want to tell. It is not about my own message.'	64.3
H	'Evidence presentations [= data visualizations] should be created in accord with the common analytical tasks at hand, which usually involve understanding causality, making multivariate comparisons, examining relevant evidence, and assessing the credibility of evidence and conclusions.' (Tufte, 2006, p.9)	64.3
I	'Complete objectivity doesn't exist, you always make choices. But it is not our aim to give our personal opinion.'	60.7
H	'Contemporary information designers seek to edify more than persuade, to exchange ideas rather than foist them on us.' (Jacobson, 1999, p.1-2)	60.7
I	'I do not give my own judgment. You have to interpret. But people have to make their own judgment.'	57.1
I	'I need to agree with the message and the visualization. But I do not show political messages.'	46.4

I	'Telling the story is the starting point. You make something with a goal. In the visualization you have to help people in the interpretation. Not simply show something. But I do not judge.'	42.9
H	'A successful visualization [...]: it informs, it makes the reader think about the world around them, and about our own lives. It stirs emotions, it encourages action, it equips us, it inspires us.' (Losowsky in Klanten, 2011, p.7)	39.3
Statements judged to express dominance of subjectivity as main goal (< 33.33%)		% objectivity
I	'You have to interpret; simply provide data is pointless. But people have to make their own judgment.'	28.6
I	'Then you must be able to experience the story. Make the story manifest. Engage people. [The question is:] how to translate a feeling, an atmosphere, into something visual?'	10.7
Statements on the characteristics of attractiveness		
I	'In interactive visualizations: playful movement, something surprising... Use of color... More feeling. Just as much information, but beautifully designed.'	
I	'Quantity matters; many data that show patterns. And then find a good form. You must be able to see information, but also an interesting form. [...]. But something that is unique and tells a story, that attracts attention and is remembered.'	
I	'For me that is in quiet, balance between text and image, about 80/20. It must contain air. It must surprise, tickle, make curious. [...] The main image must make curious. There is information everywhere, so it must have a hook in the image to which the eye lingers. [...] Quiet ordering, hierarchy in typography. Colors in a limited palette, so that you can put accents for attention.'	
I	'Exciting design, whatever that is, surprising forms.'	
I	'That is in the design. I think in metaphors, which I try to use as an illustrative element. [...] I use associations between subject and form [...]. It is mainly about information density.'	
I	'That follows from the process. It designs itself. Beauty that you see in it is a bit of intuition that you are doing alright. I don't think we have a visual language of our own.'	
I	'We start from what we consider good ourselves. That is hard to define. Data visualization is usually clear in terms of archetype. Then look at contrast, define archetype, and then the subject, and then you come in an atmosphere, and aesthetics. Aesthetics is important. We do what we like. Newspapers have guidelines, of course. Within those we search for freedom. We try to use limited colors. And silhouette: [...] try to lift the form out of the page. [...].'	
H	'Graphical elegance is often found in simplicity of design and complexity of data.' (Tufte, 2001, p.177)	
H	'Elegance is a measure of the grace and simplicity of the designed product relative to the complexity of its functions.' (Herdeg, 1981, p.8)	

Would you prefer pie or cupcakes?

Preferences for data visualization designs of professionals and laypeople in graphic design

Abstract

Data visualizations come in many different forms. In this study we investigated how professionals and laypeople in graphic design rate the attractiveness and clarity of data visualizations differing in construction type (standard or non-standard) and mode of expression (pictorial or abstract). Results showed that graphic designers rate the attractiveness of non-standard and pictorial visualizations higher than standard and abstract ones, whereas the opposite is true for laypeople. As for clarity, both groups prefer standard and abstract visualizations, which is reflected in lower response times. Results also showed that overall graphic designers' evaluations are lower than the evaluations of laypeople.

This chapter is based on:

Quispel, A. and Maes, A. (2014). Would you prefer pie or cupcakes? Preferences for data visualization designs of professionals and laypeople in graphic design. *Journal of Visual Languages & Computing*, (25)2, 107-116.

3.1 Introduction

Data visualization is a rapidly developing field within both computer science and design. Information technology is making large and complex data sets available, not only for scientists, but also for wider audiences via printed mass media and the internet. Traditionally, data visualization techniques are first of all aimed at accuracy and efficiency. But also attractiveness and aesthetics are important qualities of visualizations, especially when quantitative information has to be brought to the attention of larger audiences. Professionals in graphic design are trained to visualize messages in understandable and attractive ways. Are they able to bridge the gap between usability and aesthetics? To answer this question, we asked professionals and laypeople in graphic design to read and evaluate a selection of visualizations. The selection was a representative sample of the results of a production experiment in which graphic design professionals were asked to visualize a fictitious set of election results. That way, we collected data about production preferences of professionals, as well as data about the appreciation and efficiency of different visualization designs for both professionals and laypeople in the field of graphic design.

3.1.1 Benefits of design research

The way designers visualize information is not well documented. The design field lacks a self-definition that can support and integrate research (Storkerson, 2006). Design theorists have been struggling for decades to define their field and its position within divergent approaches toward research and theory building, without reaching consensus. Further, designers are used to working on the basis of intuition and experience, rather than explicit knowledge. As MacDonald-Ross (1977) stated 'most of the expertise in any practical art resides in people rather than on paper.' Designers, as most professional practitioners, are not used to explicitly document their methods and professional practice. They know how they solve design problems in their professional practice the same way skilled persons know how to perform their skills: it is largely tacit knowledge (Polanyi, 1966; Cross, 1982; Schön, 1983). As Schön (1983) states: 'When asked to describe their methods of inquiry, they speak of experience, trial and error, intuition, and muddling through' (p.42). Still, also according to Schön and others, there are types of research that can shed light on designers' working methods, their reasoning in action, and the resulting design choices. One of them is practice based research, e.g. examining and comparing a body of specific design cases, made in comparable situations. In the study described here we created and evaluated such a body of comparable design cases.

Designers could benefit from the insights that studies into the graphic design practice can provide, as these could enable them to move from solving one unique case after another to broader explanatory principles and solutions

for similar kinds of problems (Friedman, 2003). Scholars and practitioners involved in information visualization for broad audiences could benefit from insights into how information can be visualized in ways that are both understandable and appealing to these audiences. The experiment reported in this article attempts to contribute to these insights into the graphic design practice in that it studies the efficiency and appreciation of a comparable collection of visual displays. Although the collection was based on one straightforward set of quantitative data only, it shows a wide variety of design solutions, representing all major display formats available for visualizing quantitative information.

3.1.2 Visualizing quantitative information

Visualizations of quantitative information are ubiquitous nowadays. Since William Playfair published his first line graph in *The Commercial and Political Atlas* in 1786, we have grown accustomed to the use of bar, line and pie charts in newspapers and both popular magazines and scientific journals.

The visualization techniques that are so familiar to us now, have largely been developed by statisticians, especially during the rise of statistics in the 19th century. MacDonald-Ross (1977) wrote an excellent review of all these kinds of data visualizations and their strengths and weaknesses. These visualization techniques have been refined during the 20th century, aided by technological developments. Statisticians, computer and other scientists have elaborately described the designs of data visualizations that allow accurate and efficient readings (Bertin, 1983; Card, Mackinlay, & Shneiderman, 1999; Cleveland & McGill, 1984; Cleveland, 1985; Tufte, 1983; Spence, 2001; Ware, 2004; Ware, 2008).

In the past decades advances in computation and in graphical display software have given a strong impulse to the development of new and interactive visualization techniques (Card & MacKinlay, 1997). The term data visualization often refers to the visualization of large, complex, computer-generated data sets. The term can also be used in a broader sense and refer to the visualization of all kinds of quantitative information, from simple univariate to large multivariate data sets. In this article, we use the term data visualization in this broader sense. In our study, we used a data set of fictitious election results, with a total number of 150 elements (the number of available parliament seats) divided in nine categories (political parties).

3.1.3 Design choices and aesthetic preferences

Cognitive science has contributed much to the development of models for effective display design, based on an understanding of the way people perceive and process graphs and other external representations (Arnheim, 1969; Mackinlay, 1986; Scaife & Rogers, 1996; Carpenter & Shah, 1998; Tversky, 2011). Kosslyn's (1884) and Ware's (2004, 2008) design guidelines are based on an understanding of such perceptual and cognitive processes as well. Others

used empirical methods derived from cognitive science for testing and revising design principles (Cleveland & McGill, 1984; Heer, Kong, & Agrawala, 2009; Lloyd & Jankowski, 1999; Shah & Carpenter, 1995). Also other domains such as education show an interest in the design of visual displays (Gog & Scheiter, 2010). For an extensive overview of models for effective display design and methods for testing design principles that have been informed by cognitive science, see Hegarty (2011).

These models and guidelines all describe principles for the design of visualizations which are supposed to be clear, efficient, accurate and coming with a cognitive cost which is as low as possible. Information visualization for a broad audience however may call for a different approach, in order to grab and retain their attention, and to persuade them to retrieve the information. Perhaps other factors than accuracy and efficiency should be considered in bringing quantitative information to the attention of larger audiences, such as aesthetics.

Several theoretical models have been proposed in recent years that focus on aesthetic qualities of visual displays and aim to bridge the gap between usability and aesthetics (Judelman, 2004; Kosara, 2007; Lau & Vande Moere, 2007; Pousman, Stasko, & Mateas, 2007; Vande Moere & Purchase, 2011). When it comes to the question what aesthetic criteria exactly contribute to attractiveness, a number of empirical studies have measured the effect of a variety of design variables and attributes on user preferences, such as being abstract or pictorial (Bateman, Mandryk, Gutwin, Genest, McDine, & Brooks, 2010), 2D versus 3D (Levy, Zacks, Tversky, & Schiano, 1996); Tractinsky & Meyer, 1999), or having certain characteristics of works of art, like impressionist color palettes or abstract painting-like compositions (Fabrikant, Rebich-Hespanha, & Hegarty, 2010; Skog, Ljungblad, & Holmquist, 2003). Cawthon and Vande Moere (2007) found that perceived aesthetics was positively correlated with people's willingness to use certain data visualizations, suggesting that factors like aesthetics indeed influence the way people use visualizations. Other studies measured aesthetics in terms of subjective ratings of designs (Kurosu & Kashimura, 1995; Tractinsky, Shoval-Katz, & Ikar, 2000), or, on the other extreme, tried to capture aesthetic quality in mathematical formulae, such as metrics for characteristics such as symmetry, balance, or complexity (Berlyne, 1971; Ngo, Teo, & Byrne, 2003). All these studies reveal divergent approaches toward the notion of aesthetics. Some treat it as characteristics contributing to clarity (and, implicitly, through clarity to aesthetic experience), whereas other models treat aesthetics as design variables contributing directly to attractiveness through some sort of 'expressiveness' (Lavie & Tractinsky, 2004).

Preferences of mass media and their audiences for certain types of graph design have also been studied and are subject of an ongoing debate between designers. Zacks et al. (2002) found a preference in magazines and newspapers for the use of graphs that were 'conservative' in style; they mainly found bar charts and occasionally pie charts, colored, but rarely with background

pictures or pictographs. They also found that, despite the ease with which 3D renderings can be made of graphs with the aid of modern computer software, magazines and newspapers still publish mostly 'simple and elegant' graphs; 3D renderings were hardly used. They stated that the use of simple graphs is advocated by graphic designers, thereby referring to Tufte (1983). Tufte (actually a statistician, computer scientist and an acclaimed information designer) advocates the highest data-ink ratio in data visualization, meaning that most, if not all ink should be used to present data-information. Ink spent on other things than data-information he considers 'chartjunk', which is of no interest to the viewer. Other studies (1996) also refer to Tufte as the representative of the graphic design community, suggesting that his minimalist principles are shared by many graphic designers.

However, a look at literature that is popular among graphic designers (Klanten, 2008; Klanten, Ehmann, Bourquin, & Tissot, 2010; McCandless, 2009) and at weblogs frequented by designers engaged in data visualization (e.g. visual.ly, visualizing.org, infosthetics.com), shows a different picture. A large number of data visualizations published there seems to focus not or not only at accuracy and efficiency, but at visual pleasure as well. Likewise, Norman (2007) states that simplicity is highly overrated, and suggests that other factors should be considered as well, such as aesthetics and symbolism. Inbar, Tractinsky, and Meyer (2007) measured people's preference for standard bar graphs and minimalist versions taken from Tufte (1983). They found that people prefer non-minimalist bar graphs over the minimalist versions, but these were still simple, conventional bar charts. Bateman et al. (2010) showed that people like and remember graphics in the style of Nigel Holmes, which contain a lot of illustrative 'chartjunk', better than the plain versions in the form of simple abstract bar and line graphs. Holmes (2006), notorious for the highly illustrative information graphics he designed for *Time* magazine, claims that this visual embellishment is necessary to grab and hold the attention of not a priori interested readers.

In our study, we aim at collecting data from graphic design professionals and laypeople about two criteria or variables of aesthetics: construction type and mode of expression, of which the former is supposed to enhance ease of use (clarity, effectiveness), the latter to enhance attractiveness through expressive characteristics.

3.2 Data visualization: construction type and mode of expression

The layout of data visualizations depends, first of all, on the type of data to be represented: quantitative and geographic, quantitative and time, or quantitative and categorical data; univariate, bivariate, trivariate, or multiway data; et cetera. Certain visualization or 'mapping' techniques are more adequate for representing certain types of data or some levels of complexity of data. For example, a bar chart is apt for representing quantitative data in relation

to categorical data, whereas a line graph is more apt for representing quantitative data in relation to time (trends) (Zacks & Tversky, 1999). Still, one and the same data set can be represented by various different graphic forms, for example, both a bar and a pie chart.

Several authors have come up with taxonomies of data visualizations, often based on data types to be represented. Cleveland (1984) classified graphics as depicting one, two, or three variables. Tufte (1983) classified graphics as being relational, i.e. linking two or more variables, or not. Some classifications are based on both data type and function. Cleveland & McGill (1984) and Shneiderman (1996) categorized data visualizations according to data type and exploratory task. Other taxonomies, like the one of MacDonald-Ross (1977), are functional in nature, and focus on intended use of the diagrams. And yet others developed structural taxonomies, based on the diagrams' physical structure (Bertin, 1983; Zacks et al., 2002; Lohse, Biolsi, Walker, & Rueler, 1994).

For the purpose of this study we focused on two criteria related to the discussion above about minimal vs. less minimal design: the construction type and the mode of expression: (i) The *construction type* of a visual representation can be standard or non-standard, and (ii) the *expression mode* of a visual representation can either be abstract or it can include pictorial elements. These criteria will be explained below.

3.2.1 Construction type: standard vs. non-standard

In many situations data sets consist of a combination of categorical (nominal) and quantitative data. As Zacks et al. (2002) showed, such data are usually visualized in the form of bar and pie charts in printed mass media. A bar chart allows quick and easy comparison of the values of each category, by comparing the lengths of the bars. A pie chart allows comparing the proportions of each category to the whole, by estimating the angles of the segments. In the study reported below, we started from this standard situation, and developed a data set fitting this situation: the results of elections. The data set consisted of a total number of parliament seats ($n=150$, 100%), subdivided in the number of seats of nine political parties. We predicted that these data would be visualized mainly by these two standard construction types, bars or pies (see Appendix for examples), and based this prediction on the observation that television programs, newspapers and many other sources of election news use bar and pie charts as a standard for presenting election results. Bar charts are also classified as 'standard constructions' in the theories of Bertin (1983), one of the most influential theorists in the field of graphic design semiotics. In his view, standard constructions are the most efficient for presenting these kind of data. Bertin classified the pie chart as a 'special construction' (i.e. not the most efficient, which is the bar chart) along with donuts, stacked and divided bars, area charts and polar charts. However, recent studies (Spence & Lewandowsky, 1991) showed that pie charts can be as efficient as bar charts, depending on the task at hand. Therefore, we chose to consider both

the bar and the pie chart as the standard type, the bar chart being more apt for estimating differences among parts and the pie chart more apt for estimating proportions of parts to the whole, and both being ubiquitous in mass media nowadays in showing election results and all kinds of similar data sets.

In order to test whether bars and pies indeed represent the standard, we carried out a production experiment. We asked 41 students majoring in graphic design (19 male, 22 female) at AKV|St.Joost, Avans University, to visualize the election data described above. Each participant received a briefing on paper, instructing them to visualize the given data set in an understandable and attractive way for a broad audience. They were instructed to imagine their visualization would be published in the monthly magazine of one of the political parties (ALP) on A4 maximum, full-color. All the respondents received the same data set with only one small variation in their task: for half of the respondents, the ALP was the second largest party (28 seats), for the other half the second smallest (10 seats). They worked for about one hour on average in the classroom, individually, without cooperation and without consulting the internet. Participants were told that digital work was preferred, but they could choose to hand in sketches on paper if they wanted.

Results showed that 70% ($n = 29$) of all visualizations used bars or pies as basic design, with a dominance for bars over pies (26 vs. 3). Only 15 of these represented the data accurately. Following Tufte (2001) we consider a display accurate if (i) the spatial proportions in the display are directly proportional to the numerical quantities (proportionality condition) and (ii) if the represented portions have a common scale and origin (common scale condition). The other 14 cases used pies or bars to represent portions or somehow organized portions in a left-right or top-bottom fashion just like in a bar graph, but they either lacked a common scale (as in Appendix no. 11) or represented the proportions between parties inaccurately (as in the cupcake example, Appendix no. 12). Apart from bars and pies the graphic design students produced 12 designs based on non-standard formats (4 area, 3 polar, and 5 stacked and divided bar charts), 8 of which accurately represented the data (as in Appendix no. 19).

3.2.2 Expression mode: pictorial vs. abstract

As Zacks et al. (2002) showed, most data visualizations (bar and pie charts) that are published in printed mass media are 'simple', i.e. abstract, not containing any pictorial elements. There are however designers, like Nigel Holmes, who produce data visualizations that contain 'visual embellishment', and as Bateman et al. (2010) showed, people seem to like this kind of visualizations. As we expected that the graphic designers would not limit themselves to abstract designs, we decided to also distinguish between data visualizations in 'mode of expression' (Blackwell & Engelhardt, 2002; Engelhardt, 2002). Mode of expression refers to the extent to which graphic objects in visualizations are pictorial (ranging from highly realistic to schematic) or non-pictorial (abstract).

This difference is similar to the distinction between figurative, being high in pictorial detail, and non-figurative, being highly schematized (Richards, 2002), or iconic vs. symbolic (Tversky & Kessel, 2006). With pictorial we mean that a visualization contains graphic objects that depict recognizable physical objects or scenes.

The production experiment described above confirmed that designers to a large extent use pictorial elements in their representations. Of the 41 visualizations 25 were classified as abstract (as in Appendix, no. 2) and 16 as pictorial, with for example bars representing hats (no. 1) or a pie as a balloon (no. 8). 3 of the 25 abstract visualizations represented the data inaccurately (e.g. Appendix, no. 19); of the pictorial ones, 8 were inaccurate (e.g. Appendix, no. 20).

3.3 Evaluation study

3.3.1 Goals and expectations

As the studies above show, little is known about designers' and audiences' preferences for characteristics of data visualizations. The fact that magazines and newspapers publish mainly simple bar and pie charts does not mean that this is what their readers prefer. Perhaps it says more about budget and time constraints, as it is easy to produce simple bar and pie charts with contemporary software. The studies that compared preferences of users largely compared standard Microsoft Excel graph design options (Levy et al., 1996), or minimalist versus non-minimalist but still simple bar and pie charts (Inbar et al., 2007). But, as designers know, there are many other ways to represent data. Further, still little is known about the effect of expressive design variables such as being pictorial on the design's efficiency and about the way perceived attractiveness and efficiency interact and affect viewers' preferences.

In practice, graphs and charts for magazines and newspapers are being made by a variety of designers, like graphic designers or interaction designers, or, if simple, by journalists themselves. Graphic designers however are specifically trained to be able to convey information with visual means. They are supposed to be able to visualize ideas and information in ways that are both understandable and attractive, and to tailor their designs to the needs of their audiences. It would therefore be interesting to know to what extent graphic designers do indeed meet the needs of their audiences, and to what extent designers and laypeople share ideas about what constitutes a 'good' data visualization. Therefore, we carried out an evaluation experiment in which we asked professionals and laypeople in graphic design to evaluate data visualizations differing in construction type (standard or non-standard) and expression mode (pictorial or abstract). Also, we asked them to perform a small scale information retrieval task in order to test the speed with which they read information from these different visualization designs.

As standard types of visualizations are the types that laypeople are accustomed to, we expected laypeople to appreciate standard types more than graphic designers, who are more experienced in reading visual information. We also expected that the clarity, or efficiency, of the standard designs (the ease with which they could be read in the information retrieval task) would positively influence laypeople's overall appreciation. We expected graphic design professionals to have higher appreciation for non-standard types than for standard types, because of the relatively high number of non-standard types they produced in the production task. Further, we predicted that professionals and laypeople in graphic design appreciate pictorial visualizations more than abstract ones, based on the study by Bateman et al. (2010) and on the number of pictorial visualizations designed in the production task. As for efficiency, we expected shorter response times in the information retrieval task with standard than with non-standard types of design. Further, we aimed to find answers to the question why laypeople and designers appreciate certain types of data visualizations more than others.

3.3.2

Method

Professionals and laypeople in graphic design were asked to carry out four evaluation tasks (3 rating tasks and a selection task) and one performance task (information retrieval task).

Participants

Participants were 30 students majoring in graphic design (14 male, 16 female) at AKVSt.Joost, Avans University, who volunteered to take part in the experiment and had not participated in the production experiment, and 41 students majoring in communication and economic studies (15 male, 26 female) at Tilburg University.

Materials

We selected 20 out of the 41 visualizations produced in the production experiment (see Appendix). As to the expression mode, half of them were pictorial, half abstract. As to the construction type, 9 were standard constructions accurately representing the election data (nos. 1-9), 4 were inaccurate standard constructions (10-13), 5 accurate non-standard constructions (14-18) and two inaccurate non-standard constructions (19-20). Apart from that, the selection contained all construction types produced in the production experiment (12 bars, 2 pies, 1 stacked bar, 1 polar chart, and 4 area charts). As some of the selected visualizations were originally produced on paper, they were digitalized for the experiment. As we wanted respondents to base their appreciation and performance on visual and design aspects of the visualizations, we removed all numeric information (numbers etc.). E-prime software was used to control the random presentation of the visualizations in the

different tasks, and to collect the response times (button press) for the information retrieval task.

For the selection and explanation task, an overview of all 20 visualizations was printed on an A1 sheet of paper, randomly ordered.

Procedure

Respondents took part in the experiment individually. The experiment took about 30 minutes. Respondents were seated in front of a computer, and were instructed to carry out a number of tasks related to data visualizations. Each task was preceded by a written instruction on the screen, followed by two trials in which participants learned what buttons to use for answering the questions. After these short exercises the experimental tasks started, in the order as described below.

Attractiveness rating. In task 1, respondents were asked to rate each visualization's attractiveness. They were shown each visualization in a random order for 3 seconds. After each presentation, a new screen appeared with a five point scale (very unattractive to very attractive). Once the respondents marked one option, the next visualization appeared.

Information retrieval. In task 2, respondents were instructed that in each visualization the ALP had become either the second largest party or the second smallest party. They were asked to 'read' each visualization and answer as quickly as possible (by mouse-clicking button W for won or L for lost on the screen) whether the ALP had become the second largest (W) or the second smallest (L) party. After pressing the button, a new visualization appeared.

Clarity rating. Task 3 was the same as task 1, except for the five point scale (very unclear to very clear) and the duration of the display of each visualization: 5 seconds. Respondents were asked to rate each visualization's perceived clarity. As they performed this task after the information retrieval task, it was supposed they would base this judgment on the ease with which they had been able to retrieve the information from each visualization.

Overall rating. Task 4 was similar to 1 and 3: the respondents had to give an overall mark on a 10-point scale (extremely bad - extremely good). They were instructed that this mark reflected their opinion about the overall quality of the visualization, all things considered. Visualizations appeared in a random order one by one. Visualizations were displayed until participants marked them on a ten point scale presented below on the screen, after which the next visualization appeared.

After that, the respondents were asked to sit at another table where all visualizations were presented together on one A1 sheet.

Selection task. In task 5 respondents were presented with all visualizations on one A1 page; they were asked to select the three designs they appreciated most, all things considered, and the three they appreciated least. Afterwards they were asked to explain their selection. Responses were audio recorded.

Data analysis

For the rating tasks (tasks 1, 3 and 4) as well as for the information retrieval task (task 2), the data were aggregated by construction type, expression mode and participant. Means were compared using univariate analysis of variance. Response times higher than two standard deviations from the mean were considered outliers and were left out of the analysis.

As we did not provide respondents with numerical information, nor with prior information about the election results, we were not interested in the correctness of their information retrieval task. Responses showed no effects of type, mode or design experience. For almost all items the answer (won or lost) was correct in 80 to 100% of the cases, with three notorious exceptions: nos. 12 and 13 which did not show any information about proportions, and one deceiving one (no. 17), that placed the losing party at the top of a pyramid.

As for construction type, we compared two groups of items: on the one hand, all items with a standard construction type (bar/pie, nos. 1-13); on the other hand the items using a non-standard design (nos. 14-20). That way, each group consists of visualizations which accurately represent the data (standard: nos. 1-9, non-standard: nos. 14-18) and inaccurate visualizations (standard: nos. 10-13; non-standard: nos. 19-20). We did not expect accuracy to play a major role, as respondents were unaware of the exact proportions of the election results. This expectation was confirmed when we repeated the analyses below leaving out all inaccurate items. This analysis showed the same effects as the analyses reported below.

For the selection task (task 5), the responses of each participant were listed. Participants were allowed to give more than one reason for selecting a visualization, e.g. *it is clear and attractive*. This resulted in a list containing at least three reasons (for three selected visualizations) for the '*best of*' selection and at least three reasons for the '*worst of*' selection per participant. The explanations for the *best of* selection were clustered into three categories of reasons: Clear (e.g., *it is clear, easy to read, you see the differences at a glance*), Attractive (e.g., *it is attractive, funny, beautiful, nicely looking*), Different (e.g., *it is unusual, different, not standard, unconventional*). Likewise, the explanations for the *worst of* selection were clustered in two categories: Unclear (e.g. *it is not clear, it is very unclear, I can't see what it is about, I don't understand it, there is no information*) and Unattractive (e.g., *it is ugly, unattractive*). When participants gave more than one reason, each one was counted. Only a few infrequent comments could not be classified in one of these categories (n = 19; 3.5%). They were disregarded.

The results of this explanation part were analyzed using an independent-samples T test.

3.3.3 Results

Rating tasks and information retrieval task

Table 1 shows the results for the rating tasks and the response times in the information retrieval task. Results are reported separately for type and mode.

			professionals	laypeople
type	attractiveness	standard	2.13 (.66)	3.02 (.67)
		non-standard	2.76 (.68)	2.26 (.68)
	clarity	standard	3.02 (.84)	3.19 (.86)
		non-standard	2.42 (.76)	2.46 (.66)
	overall grade	standard	4.84 (1.09)	5.84 (1.28)
		non-standard	5.17 (1.24)	4.75 (1.19)
response time	standard	3153 (1276)	3020 (1070)	
	non-standard	4428 (1896)	4458 (1644)	
			professionals	laypeople
mode	attractiveness	abstract	2.23 (.76)	2.77 (.75)
		pictorial	2.65 (.67)	2.51 (.78)
	clarity	abstract	3.08 (.83)	3.20 (.83)
		pictorial	2.36 (.72)	2.45 (.68)
	overall grade	abstract	5.09 (1.09)	5.79 (1.31)
		pictorial	4.92 (1.26)	4.80 (1.20)
	response time	abstract	3305 (1371)	3171 (1192)
		pictorial	4273 (1922)	4306 (1680)

Table 1 Attractiveness (1-5), clarity (1-5) overall grade (1-10), and response time (ms) of visualization categories (type: standard - non-standard; mode: abstract - pictorial) related to design experience (professionals - laypeople). Means (standard deviations between brackets).

Construction type

Attractiveness rating. There was a main effect of design experience ($F(1, 280) = 5.74, p < .05$); laypeople gave higher attractiveness ratings than design professionals. There was a marginal effect of type ($F(1, 280) = 3.51, p = .06$) and a significant interaction between design experience and type ($F(1, 280) = 73.36, p < .001$). Based on the latter, we performed a split analysis for both groups. This analysis showed that both for the design professionals and for the laypeople there was a significant effect of type on attractiveness rating (professionals: $F(1, 118) = 26.30, p < .001$; laypeople: $F(1, 162) = 51.57, p < .001$). As expected, the design professionals rated non-standard visualizations higher than standard ones, whereas laypeople rated standard visualizations higher than non-standard ones.

Clarity rating. There was a main effect of type on clarity rating ($F(1, 280) = 49.96$, $p < .001$), with standard types being rated higher than non-standard ones by both design professionals and laypeople. There was no effect of design experience ($F(1, 280) = 1.15$, $p = .28$), neither an interaction between design experience and type ($F < 1$).

Overall rating. The analysis of variance showed a main effect of design experience ($F(1, 280) = 4.04$, $p < .05$) and type ($F(1, 280) = 6.71$, $p < .05$) on overall rating. Laypeople gave higher overall grades than design professionals. Also, there was a significant interaction between design experience and type ($F(1, 280) = 23.96$, $p < .001$). A split analysis showed a significant effect of type on overall rating for the laypeople ($F(1, 162) = 31.63$, $p < .001$). Laypeople rated standard types higher than non-standard types.

Response times. There was a main effect of type on response times ($F(1, 279) = 57.45$, $p < .001$). Response times were higher for non-standard than for standard types. There was no effect of design experience ($F < 1$), neither an interaction between design experience and type ($F < 1$).

Expression mode

Attractiveness rating. There was a main effect of design experience ($F(1, 280) = 4.73$, $p < .05$). Laypeople gave higher ratings than design professionals. There was no effect of mode ($F < 1$), but there was a significant interaction between mode and design experience ($F(1, 280) = 14.53$, $p < .001$). Based on the latter, a split analysis was performed, which showed an effect of mode on attractiveness rating for both design professionals and laypeople (professionals: $F(1, 118) = 10.54$, $p < .05$); laypeople: $F(1, 162) = 4.66$, $p < .05$). Design professionals rated pictorial visualizations higher than abstract ones, whereas laypeople rated abstract visualizations higher than pictorial ones.

Clarity rating. There was a main effect of mode on clarity rating ($F(1, 280) = 63.53$, $p < .001$). Abstract visualizations were rated higher than pictorial ones. There was no effect of design experience ($F(1, 280) = 1.19$, $p = .28$) and no interaction between design experience and mode ($F < 1$).

Overall rating. The analysis showed a main effect of design experience on the overall rating ($F(1, 280) = 3.94$, $p < .05$). Designers gave lower ratings than laypeople. There was also a main effect of mode on the overall rating ($F(1, 280) = 15.54$, $p < .001$) and a significant interaction between mode and design experience ($F(1, 280) = 7.84$, $p < .05$). A split analysis showed an effect of mode on overall rating for the laypeople ($F(1, 162) = 25.50$, $p < .001$): laypeople rated abstract visualizations higher than pictorial ones.

Response times. There was a main effect of mode on response times ($F(1,279) = 31.81, p < .001$). Response times were higher for pictorial visualizations than for abstract ones. There was no effect of design experience ($F < 1$) and no interaction between the two factors ($F < 1$).

Selection task and explanation

Selection task. The selection task resulted in a list of most and least appreciated visualizations by graphic design professionals and laypeople. Table 2 presents the top 5 of most and least appreciated designs. This top 5 represents the 5 visualizations that were mentioned most often as being one of the three best and one of the three worst of all twenty visualizations.

	Professionals		Laypeople		Overlap	
	<i>n</i>	<i>item number</i>	<i>n</i>	<i>item number</i>	<i>n</i>	<i>item number</i>
best	5	1, 11, 15, 16, 19	5	1, 2, 3, 5, 6	1	1
standard	1		5		1	
pictorial	3		1		1	
worst	5	4, 9, 12, 13, 20	5	12, 13, 15, 17, 20	3	12, 13, 20
standard	1		0		0	
pictorial	4		4		3	

Table 2 Top 5 most and least appreciated visualizations by graphic design professionals and laypeople

As expected, laypeople appreciate standard types of construction more than deviating types: all visualizations chosen as best are standard. The graphic design professionals, on the other hand, appreciate non-standard types more (4 out of 5). Also as expected, the professionals seem to appreciate pictorial types more than abstract types: the majority (3 out of 5) is pictorial. On the other hand, the laypeople chose only one pictorial type, suggesting they appreciate abstract types more than pictorial ones. The two groups have only one preference for a visualization in common, namely for a standard bar chart adding a little bit of pictorial fun.

The *worst of* selection task shows a much higher degree of overlap between laypeople and professionals. In their dislikes, the laypeople and the designers agree on three visualizations (Appendix, nos. 12, 13, and 20). All three are deviating in that proportionality is distorted; two of them, a row of cupcakes and a series of chairs (nos. 12 and 13) don't show any differences in proportions at all.

Explanation task. Designers and laypeople differ significantly in the reasons they give for their choices, as is shown in table 3 below. When asked to explain why they appreciate certain visualizations more than the others, professionals mention attractiveness more often than laypeople ($t(69) = 3.19, p = .002$),

whereas laypeople mention clarity more often than professionals ($t(69) = 4.06$, $p < .001$). Further, professionals tend to mention 'being different' more often than laypeople as a reason for appreciation ($t(69) = 2.06$, $p < .05$).

Also in the reasons they give for their choices of least appreciated visualizations professionals and laypeople differ significantly. Being unclear is mentioned more often by laypeople ($t(69) = 3.35$, $p = .001$), whereas unattractiveness is more often a reason for dislike for professionals ($t(69) = 3.59$, $p = .001$).

	Professionals M (SD)	Laypeople M (SD)
Best		
Clear	0.98 (1.19)	2.10 (1.09)
Attractive	2.73 (0.63)	2.07 (1.11)
Different	0.77 (1.07)	0.34 (0.66)
Worst		
Unclear	2.13 (1.07)	2.78 (0.53)
Unattractive	1.07 (1.05)	0.37 (0.58)

Table 3 Reasons mentioned for 'best of' and 'worst of' selection related to design experience (professionals, laypeople). Means per participant (standard deviations between brackets).

3.4 Discussion

The results show clear differences between the two target groups in the study: professionals rate the attractiveness of non-standard and pictorial visualizations higher than standard and abstract versions. Laypeople prefer standard and abstract visualizations. The clarity ratings do not follow the same pattern: standard and abstract visualizations are preferred for both target groups. For laypeople, the overall ratings of visualizations are in line with their attractiveness ratings, with higher ratings for standard and abstract visualizations. For professionals, there is no significant difference between types and modes. The response times for the two groups are in line with their clarity ratings: longer response times for the non-standard and pictorial visualizations. As expected, design professionals show a clear preference for non-standard types of visualizations, whereas laypeople prefer standard types.

These results largely follow our expectations, except on one point: laypeople do not appreciate pictorial visualizations more than abstract ones, as we expected based on the results found by Bateman et al. (2010). In their study the pictorial visualizations were colorful, whereas the abstract versions were very plain, black and white graphs, which may have influenced participants' preferences for pictorial versions. In our study, they actually preferred the standard and abstract types of visualizations as they are usually published in mass media. Still, one standard and pictorial design is among the most

appreciated ones, both in the group of laypeople and in the group of graphic design professionals. This suggests that there may be a type of design, both standard (easily readable) and pictorial, that both the laypeople and the graphic design professionals appreciate.

Looking at the designs that both groups chose as worst, two designs stand out for the fact that both designers and laypeople think that these are bad visualizations and they agree on the reason why: these designs show no information about proportions at all. One only shows cupcakes, one differing in color, which seems to indicate that one is a winner. The other shows only two rows of chairs (seats), also with one differing in color. Both designs are nondescript in terms of proportionality: they don't show any differences in proportions at all.

The fact that the majority of the types least appreciated by the designers is pictorial, may be caused by the fact that these two non-informative designs happened to be pictorial types. Also, the fact that response times were longer for pictorial types, may be caused by the fact that these made up for the majority of the visualizations that were disproportional or lacked a common scale (5 of 6). Apparently, disproportional and/or lack of a common scale cause more interpretation difficulty.

Both groups differ also clearly in the reasons they give for their preferences. The laypeople put more emphasis on clarity, whereas the design professionals attach more value to attractiveness. The fact that laypeople put more emphasis on clarity may account at least in part for their preference for standard types. After all, standard types are by definition the most efficient, easiest to read. This is confirmed by the fact that response times in the information retrieval task were higher for non-standard types than for standard types.

In all, the results show that there is a clear difference in preferences for design types between graphic design professionals and laypeople in graphic design. Especially the difference in preference for standard and non-standard types of visualizations raises questions about the extent to which graphic designers can indeed bridge the gap between usability and aesthetics in data visualization. The design professionals don't value clarity that much, they value attractiveness more. If it is among designers' tasks to tailor designs to the needs of their audiences, this means they would do well to make sure they test their designs before publishing. It is common among designers to 'test' designs in an informal way, often among fellow designers in the design company, or friendly colleagues. However, testing designs among a group of laypeople in the field of graphic design might yield valuable insights into the way their designs are appreciated by their audiences.

There are some limitations to the study reported on. The preferences in this study were studied using one specific set of data, election results, so generalizations should be made with caution. On the other hand, the fact that election results are such a common kind of data to be visualized in mass media, and the visualizations used in the evaluation study show such a wide variety of

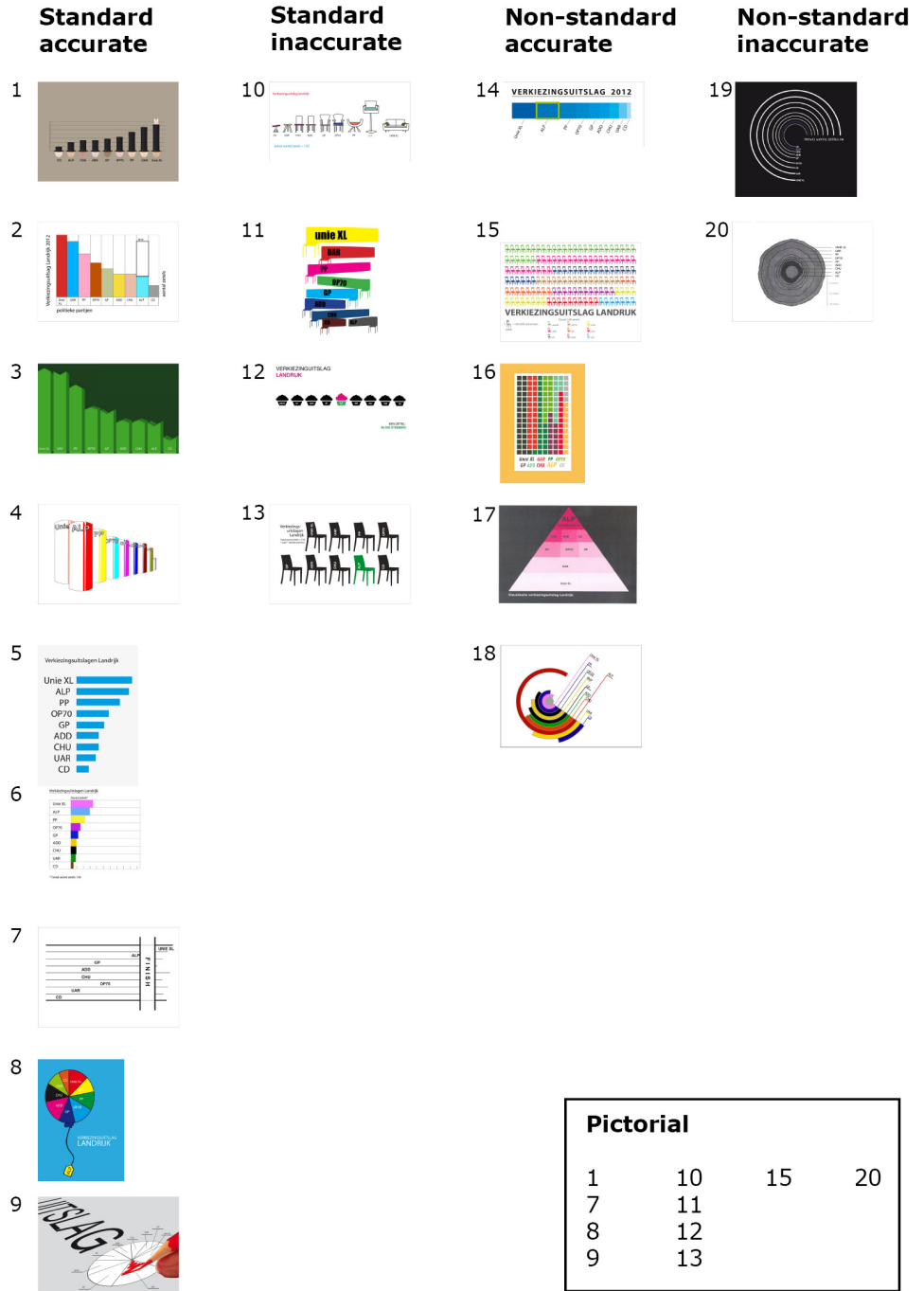
designs, give reason to believe that similar results would be found with other, similar (category x quantity) data sets.

Further, materials and tasks used in this experiment did not enable us to unambiguously determine the effect of non-standard construction type and inaccuracy in data presentation. These variables can better be tested using constructed materials in which these two variables are varied more systematically, by using a better balance between evaluation and retrieval tasks and by providing respondents with prior knowledge about the data represented in the visualizations.

Would you prefer pie or cupcakes?

Appendix

Visualizations used in the experiment, classified according to type of construction and mode of expression.



Graph and chart aesthetics for experts and laypeople in design: The role of familiarity and perceived ease of use

Abstract

We investigated the relationship between familiarity, perceived ease of use, and attractiveness of graph designs in two target groups: experts and laypeople in design. In a first study we presented them with a variety of more and less common graph designs, and asked them without any additional task to evaluate their familiarity, attractiveness, and perceived ease of use. They judged the familiarity and ease of use of the graphs similarly, but they differed in their attractiveness judgments. Familiarity and perceived ease of use appeared to predict attractiveness, but stronger for laypeople than for designers. Laypeople are attracted to designs they perceive as familiar and easy to use. Designers are attracted to designs between familiar and novel. In a second study we asked designers and laypeople to first perform an information retrieval task with the same graphs, and then rate their attractiveness. Laypeople's appreciations remained the same, but the designers' judgments of attractiveness were different from those in study 1. Correlational analyses suggest that their attractiveness judgments after use were affected not by actual usability, but by perceived ease of use of the graphs.

This chapter is based on:

Quispel, A., Maes, A., and Schilperoord, J. (in press, first published online September 2015).
Graph and chart aesthetics for experts and laymen in design: The role of familiarity and perceived ease of use. *Information Visualization*.

4.1 Graph aesthetics

Historically, data visualizations were primarily meant for use by statisticians and scientists. Their designs were aimed at an efficient and accurate reading of the data in analytical tasks. But data visualizations are growing popular. Bringing quantitative information to the attention of larger audiences may call for other qualities than accuracy and efficiency alone, such as attractiveness. This study addresses the aesthetic appeal of data visualizations, or graphs, which are not only becoming increasingly popular, but also show a growing variation in designs in everyday mass media. We investigated the aesthetic judgment of graphs within two target groups: graphic designers, as producers of data visualizations, and laypeople in design, as their audience. Graphic designers are trained to tailor their designs to the needs of their audiences. Therefore, the main question we address in this paper is the extent to which the two groups share opinions with regard to the aesthetics of graphs.

The term aesthetics has had several connotations through the centuries. Originally it referred to the study of sensory perceptions, but since the 18th century it is commonly conceived as the study of beauty and fine art. Researchers studying displays like graphs use many terms to refer to the aesthetic value of these displays: sensory pleasantness (Hekkert & Leder, 2008), beauty (Tractinsky, 2004), attractiveness (Vande Moere & Purchase, 2011). All of these terms refer to a similar aspect of displays. But whereas the term 'beauty' seems more appropriate for visual 'displays' that are not directly intended for use (like works of art or natural sceneries), the terms 'visual pleasantness' and 'attractiveness' seem more appropriate for evaluating functional visualizations as the ones we focus on in our studies. Therefore, we prefer to use 'attractiveness' to refer to the main topic of this paper.

With the growing interest in data visualization and its aesthetics in design and artistic practices, researchers in the information visualization community (specialized in visualization of large, complex data sets) also have come to recognize and propose aesthetics and the interplay between aesthetics and usability as an important research subject (Chen, 2005). Several theoretical models have been proposed to bridge the gap between usability and aesthetics (Burkhard, Andrienko, Andrienko, et al., 2007; Judelman, 2004; Kosara, 2007; Lau & Vande Moere, 2007). These models share the viewpoint that the visualization community and the artistic community could benefit from each other's knowledge on functional and visual qualities, but they shed little light on which variables actually contribute to aesthetic experience, and how. On the other hand, several empirical studies attempting to reveal such aesthetic factors have been conducted, and show divergent approaches toward the notion of aesthetics.

Some studies attempt to reveal objective perceptual variables affecting aesthetic experiences. For example, the difference in preferences has been studied between 'embellished', pictorial graphs and plain abstract graphs (Bateman, Mandryk, Gutwin, et al., 2010). It appeared that participants'

descriptions of the information in the two graph types was equally accurate, but that pictorial, embellished types were recalled better. Preferences for 2D vs. 3D bar graphs have been studied (Levy, Zacks, Tversky, & Schiano, 1996), showing that people prefer 2D versions for immediate, analytical use, but choose to use 3D versions when they have to show quantitative data to others and want them to be remembered. Some studies have also investigated the influence of modern painting like color palettes on appreciation of topographic maps (Fabrikant, Christophe, Papastefanou, et al., 2012). Rating and ranking tasks and skin responses showed that participants' emotional responses can be systematically measured and analyzed to study the effect of aesthetic criteria. In another line of research studies are conducted that aim to develop ways to attract viewers' attention and to engage them in visualizations by applying aesthetically pleasing non-photorealistic rendering styles to data visualizations (Rheingans & Ebert, 2001; Healey, Tateosian, Enns, & Remple, 2004).

Other studies are based on the assumption that people are attracted to clearness. In these studies aesthetics is treated as a characteristic contributing to clarity, and through clarity to visual appeal. In models within computer sciences, for example, aesthetic quality is defined as a set of quantifiable metrics. These models use algorithms to measure features such as order, balance, symmetry, or complexity (Ngo, Teo, & Byrne, 2003; Ware, Purchase, Colpoys, et al., 2002). In information visualization the term 'graph aesthetics' commonly refers to heuristics for enhancing the readability of node-link diagrams, for example by minimizing the number of bends.

The influence of such graphical variables on aesthetic perceptions is measured in several ways. Some studies capture aesthetic experiences objectively, measuring skin reactions associated with pleasant feelings (Fabrikant et al., 2012) or activity in brain areas associated with rewarding feelings using MRI (Aharon, Etcoff, Ariely, et al., 2001). Most studies in the fields of data visualization, product design and HCI use subjective measures of preference or liking. In these studies aesthetic perceptions are measured by asking participants to rate the beauty or attractiveness of visualizations (Ngo et al., 2003; Cawthon & Vande Moere, 2007; Hekkert, Snelders, & Van Wieringen, 2003; Tractinsky & Meyer, 1999), to rank visualizations from least to most preferred (Cawthon & Vande Moere, 2007), or to choose one of several designs for use in a specific situation (Levy et al., 1996; Tractinsky & Meyer, 1999).

These studies conjecture that aesthetics is to be found first and foremost in properties of the object itself. At the same time, there is growing awareness that people's aesthetic evaluation of graphs may be influenced by the way in which they interact with graphs. In this study we define aesthetics as the subjective evaluation of a graph's attractiveness. We do not attempt to reveal properties of graphs affecting their attractiveness. Instead, we investigate how people's experience with graphs may affect their attractiveness. In the remainder of this article we will discuss the role of three variables potentially affecting the aesthetic evaluation of graph designs: familiarity, perceived ease

of use, and actual use. In particular, we will focus on how these variables differently affect judgments of experts and laypeople in design. After that, we present two empirical evaluation studies. In the first study we asked experts and laypeople in design to evaluate the familiarity, perceived ease of use, and attractiveness of a set of more and less conventional graph designs. In the second study we investigated the effect of actual usage of the graphs in an information retrieval task on evaluations of attractiveness by both groups. We conclude with a discussion of the results.

4.1.1 Familiarity and aesthetics

Ancient and modern theories about the nature of beauty make different predictions about the relationship between prior exposure or familiarity and aesthetics or attractiveness.

Classical theories about the nature of beauty advocate a balance between extremes in order to achieve appreciation and attention, for example between chaos and order, expectation and surprise (Berlyne, 1971). In ancient Greece, Plato (in the *Statesman*) defined beauty as 'a standard removed from the extremes', and according to Aristotle 'a master of any art seeks the intermediate' (Berlyne, 1971). Also during the Renaissance beauty was believed to be found in an equilibrium between mutually counterbalancing factors (Descartes, 1650; Hutcheson, 1725). Novelty, originality, and variety were supposed to be necessary to make things lively, whereas familiarity, coherence and economy would be necessary to prevent chaos. The idea of balance between extremes can also be found in more recent models where aesthetic preferences are described in terms of a balance between novelty and familiarity. For example, evolutionary aesthetics (Hekkert, 2006) proposes that human beings derive aesthetic pleasure from phenomena that help them survive. On the one hand they are attracted to familiar things, because familiarity, as a result of repeated exposure, facilitates perceptual organization and helps them to bring order in a complex world (Hekkert, 2006; Reber, Schwartz, & Winkielman, 2004). On the other hand people are also attracted to new, unusual things, presumably because novelty facilitates learning, which is also a vital capacity for survival (Hekkert, 2006). There is some experimental evidence for the assumption that aesthetic pleasure can be found in a balance between familiarity and novelty. Within experimental psychology, Berlyne (1971) carried out several empirical studies to test the balance-between-extremes theories, including the balance between familiarity and novelty. Although his studies showed some evidence of the influence of both familiarity and novelty on aesthetic pleasure, several studies remained inconclusive and some results also have been contradicted by other studies (Zajonc, 1968, 1984). Studies in product design have shown that people indeed prefer product designs that balance typicality (a notion related to familiarity) and novelty (Hekkert et al., 2003). Participants were asked to evaluate a series of designs in terms of good (typical for the product category, e.g. 'teapots') or poor examples

of the category in question, and in terms of aesthetic appeal. These studies showed that people prefer novel designs as long as the novelty does not affect typicality. In other words, a teapot that is unlike other teapots is nice, but only if it still resembles a teapot.

Other theories however predict that attractiveness linearly increases with increasing familiarity. For example, according to Zajonc (1968, 1984) it is merely repeated exposure to a stimulus that increases its aesthetic appreciation. According to Reber et al. (2004) beauty is grounded in processing experiences. They propose that repeated exposure to a stimulus (which leads to familiarity) makes perceptual and cognitive processes easier and more fluent, and that this fluency is perceived as pleasing. Familiarity could also be aesthetically pleasing because it signals that a stimulus is probably not harmful (Zajonc, 1968, 1984). Dislike of novel stimuli could be a precognitive biological mechanism that makes people cautious in the case of potentially harmful stimuli. Several empirical studies have also supported this proposition that the more fluently someone can process an object, the more positive his or her aesthetic judgment is (for a review of such studies, see Reber et al., 2004) This would suggest that people generally prefer familiar visualizations.

How can we apply these theories to the difference between experts and laypeople in design, the two target groups we asked to rate the familiarity and attractiveness of a series of more or less conventional and novel graph designs? Familiarity is supposed to be the result of previous exposure. Design experts are accustomed to looking at and working with a variety of visual displays on a daily basis. Therefore, we expected that overall, they will perceive all graph designs as more familiar than non-designers. However, as the relative frequency of each graph should be the same for both groups, we expected that the rank ordering of the various graphs on the familiarity scale will be similar.

As for the relationship between familiarity and attractiveness, classical theories predict that both groups find designs in between familiar and novel the most attractive, whereas processing fluency theories predict a linear positive relation between familiarity and attractiveness. But there are reasons to expect differences in the evaluations of attractiveness between the two groups. Studies into aesthetic experiences of art works have shown differences between experts (with art training) and novices (laypeople) in aesthetic preferences, with novices preferring simple and prototypical stimuli and experts preferring complex and novel stimuli (McWhinnie, 1968). This is also suggested by Reber et al. (2004): they propose that aesthetic pleasure increases with increasing processing fluency, but acknowledge that in some cases people appear to be attracted to more complex and novel stimuli. It has been suggested that these differences may be caused by the fact that experts tend to consider not only the aesthetic experience itself (the pleasing of the senses), but also the aesthetic value of stimuli, in terms of what they consider good or bad taste (Bourdieu, 1987; Gombrich, 1995). As a result, they may evaluate familiar stimuli more negatively than novel ones, despite the high

processing fluency. Although these studies focused specifically on the appreciation of works of art, we assume that similar mechanisms apply in design. Designers are well trained in processing visual information and tend to attach more value to visual appeal than to functionality (Quispel & Maes, 2014). They may therefore consider familiar designs less appealing because of a lack of originality and visual challenge. For this reason, it may be expected that designers are more attracted to novel data visualizations than laypeople, and that the latter appreciate familiar designs.

4.1.2 Familiarity, attractiveness, and perceived ease of use

Several studies have shown relationships between aesthetics and perceived ease of use. In these studies a distinction is made between perceived ease of use of displays based on visual impressions only, and perceived ease of use after displays have actually been used to carry out a certain task. Most of the studies describing relationships between aesthetics and (perceived) usability are situated in the field of human-computer interaction (HCI). In HCI, usability not only involves understanding and interpretation, but also interaction with the design through a computer interface. In our study, we evaluated the perceived ease of use of static graphs without use, based on visual impressions, and after use in an information retrieval task. Because carrying out a task with a static graph does not imply interaction, but only interpretation of its informational content, we measured ease of use by assessing ease of interpretation.

The first study in which a relationship was shown between aesthetics and usability is Kurosu and Kashimura's (1995). They asked participants to rate the perceived ease of use and the beauty of several user interfaces (without having to use them in functional tasks) and found that perceived ease of use was positively correlated with perceived beauty. The authors then calculated the 'inherent usability' of the interfaces, based on design characteristics believed by interface designers to enhance usability, such as certain arrangements and groupings of keys. They only found a significant correlation between perceived ease of use and the usability measure of familiarity.

Tractinsky, Katz, and Ikar (2000) tested the relationship between aesthetics and perceived ease of use before and after actual use of ATM layouts. They asked participants to rate the layouts' aesthetics and perceived usability before use and found high correlations between aesthetics and perceived usability. Then participants performed tasks with the layouts, while usability – task completion times – was manipulated with dysfunctional buttons that caused delays. Results showed that perceptions of aesthetics and usability were not affected by these usability manipulations if participants worked with layouts that they already found highly aesthetic before use. They perceived the interfaces as equally aesthetic and equally usable as before use.

An explanation for the relationship between these perceptions of aesthetics and usability may be found in the concept of familiarity. As we described in

the previous section, processing fluency theories submit that familiar things are found attractive, because familiar things are perceived as easy to use (see Figure 1). Tractinsky et al. (2000) did not measure the familiarity of the displays they evaluated. Perhaps participants in their study considered designs more attractive that were also more familiar to them and, as a result of this, were perceived as easy to use. This explanation is supported by the high correlation found between perceived ease of use and familiarity in Kurosu and Kashimura's study (1995).

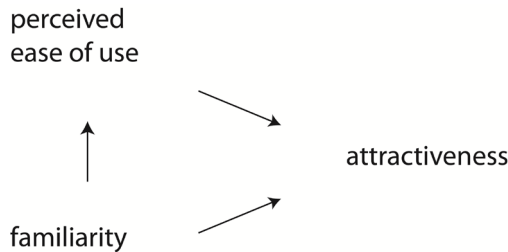


Figure 1 Proposed relationships between familiarity, perceived ease of use, and attractiveness

In our studies, we started from the strong link between familiarity and ease of use as found in the literature and studies described above. The more familiar devices are, the more they are perceived as easy to use. Therefore, we expected strong correlations between familiarity and perceived ease of use, both for designers and laypeople. However, the focus in our studies concerns the way familiarity and perceived ease of use are associated with evaluations of attractiveness. Here, we expect to find differences between laypeople and designers. Based on processing fluency theories and Tractinsky et al.'s study (2000), we expected familiarity and perceived ease of use to predict a large amount of the attractiveness values for the laypeople. Based on studies showing differences between novices and experts (McWhinnie, 1968), and a previous study into preferences of designers and laypeople (Quispel & Maes, 2014), we expected familiarity and ease of use to be stronger predictors of attractiveness for laypeople than for designers. Also based on studies showing differences in preferences between experts and novices, we expected designers to prefer more novel designs than laypeople. Therefore, we expected the relationship between familiarity and attractiveness to show in an inverted U-shape for the designers (designs between familiar and novel being found the most attractive). For laypeople classical theories give reason to expect an inverted U-shaped relationship, whereas processing fluency theories predict a linear relationship.

Finally, we wondered if actual use would affect perceptions of attractiveness, and if differences would occur here between designers and laypeople. The results in Tractinsky et al.'s study showed that perceptions of aesthetics and usability were not affected by use, even if using the devices was frustrated (by adding dysfunctional buttons). Therefore, we expected that fairly simple,

common tasks with the graphs (as we asked the respondents to perform in study 2) would not change laypeople's and designers' judgments.

4.1.3 Research goals

Our main goal in this study was to investigate, for design experts and laypeople in design, the relationships between attractiveness, familiarity, and perceived ease of use. On the basis of the considerations given above, we had the following expectations.

- 1 We expected designers to be more familiar with all design types than laypeople in design.
- 2 We expected rankings of familiarity ratings to be similar for both groups.
- 3 Correlations between familiarity and perceived ease of use were expected to be significant for both groups.
- 4a We expected familiarity and perceived ease of use to be predictors of attractiveness for both groups. The two factors were expected to be stronger predictors for laypeople than for experts.
- 4b For laypeople, we expected either a linear or an inverted U-shape relation between familiarity and attractiveness. For designers we expected an inverted U-shape relation.
- 5 We expected actual use of the graphs not to affect judgments of attractiveness.

We address the first four hypotheses in study 1, where we asked participants to rate the familiarity, perceived ease of use, and attractiveness of a set of graph designs. The last hypothesis is addressed in study 2, an explorative study in which we asked participants to perform an information retrieval task with the graphs and then to rate their attractiveness.

4.2 Evaluation study 1

We asked designers and laypeople to rate the familiarity, attractiveness, and perceived ease of interpretation of 12 more and less conventional graphs in an online survey. These variables were measured using a 7-point Likert scale.

4.2.1 Method

Participants. 272 laypeople in design (79 women) and 44 design experts (28 women) participated in the survey. Laypeople's data were collected using the crowdsourcing service CrowdFlower (www.crowdfunder.com), a platform of which an evaluation study has shown that it is suitable to obtain high-quality data (Buhrmester, Kwang, & Gosling, 2011), especially for relatively easy tasks such as a rating task. CrowdFlower uses a built-in quality control system banning respondents who have shown to yield unreliable results. The survey was taken by participants in 15 western countries (for \$0.30 compensation).

The results of 23 respondents were removed because they had filled in exactly the same ratings for each question, or because they indicated to have design experience or education. Design expert data were collected by sending an email with an invitation to take part in the survey to students at the departments of graphic and visual communication design in Cambridge (UK) and New York (US). Participation was voluntary. The majority of the laypeople (62%) and the designers (80%) were between the ages of 18 and 34. Of the laypeople 58.7% followed vocational (19.7%) or university/polytechnic (39%) education, and the others high school (41.3%) as their highest educational level. All of the designers followed university/polytechnic studies. In the analyses we only included the results from laypeople with educational levels similar to those of the designers (university/polytechnic, n=97, 30 women).

Materials. We designed a set of 12 static graphs (see Figure 2). The graphs represented a combination of quantitative and nominal data, thus mirroring the majority of graphs in newspapers and magazines (think of election results, or budget cuts). The 12 graph designs were chosen to reflect the diversity of visualization techniques with which a combination of quantitative and nominal data can be represented. The construction of a graph is determined by the dimensions of the plane ('surface') used to represent quantity (e.g. length or area), and by the particular way in which these dimensions are portrayed (e.g. rectangular or circular). The test materials contained both graphs using length and graphs using area to represent quantity, and both circular and rectangular forms. We decided not to use line graph designs, since these encode different types of data (ordinal or interval data), and they do not offer much variation in design.

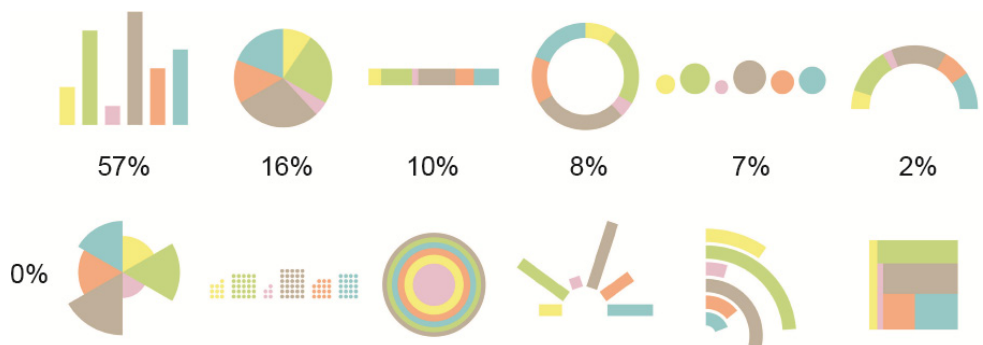


Figure 2. Tested graphs and their proportions in the pilot study count

In a pilot study, we checked the distribution of conventional and new designs by counting the frequency of the graph types in everyday mass media. These counts mirror the level of exposure of people to each of the graphs, and thus guarantee a broad distribution of the familiarity ratings in our data. Over a period of six weeks (April-May 2014), we counted 136 graphs representing nominal x quantitative data in three Dutch national newspapers (Volkskrant,

NRC, Telegraaf) and three opinion magazines (Time, Elsevier, Groene Amsterdammer). As shown in Figure 2, six of the graph designs were used in the media, some frequently and some rarely (57% bar graphs, 16% pie graphs, 10% divided bars, 8% donut charts, 7% bubble charts, and 2% semi-circles); the other six were not found in the pilot study at all.

The 12 graphs were all designed using an identical color palette, and each graph represented the same six proportions. So, they all represented the same data set, but the graphs were not related to a specific functional context enabling participants to interpret the data. No numbers or scales were included, to make sure that judgments were based only on the visual structure of the graphs. The pictures of the graphs measured 500 pixels (width) x 510 pixels (height). As we used an online survey we had no control over the sizes of the screens on which the survey was taken.

To measure attractiveness, familiarity, and perceived ease of use, 7-point Likert scales were used. The attribute familiarity was chosen to capture the degree of exposure of respondents to the graphs. We opted for attractive as an adequate generic option to elicit an aesthetic judgment about functional visualizations as the ones we presented to them. Finally, as explained above, we opted to inquire about ease of interpretation to gauge their judgment about the perceived ease of use of the graphs. We realize that single-item measures may be less valid than multi-item construct measures. Still, we considered possible disadvantages less important than the danger that repeated exposure to too many questions could result in overloading respondents.

Procedure. The survey started with written instructions. The instruction explained that participants were about to look at graphs as they could appear in a journal or magazine or on a website, and that each graph could be used to represent the same kind of data, for example election results per political party, or budget cuts per public sector. The instructions were followed by a practice example, after which the respondents went through 12 screens with the graphs (presented in a random order) at their own pace (by pressing the forward button). Each screen presented a graph centered at the top without any legends or text labels. At the bottom of the screen the same three 7-point Likert scales were presented in the same order (see Figure 3).

The scales ranged from 'strongly disagree' to 'strongly agree'. Participants were asked to indicate to what extent they thought each graph was familiar, attractive, and easy to interpret. The survey could only be completed when each question was answered. After the rating task, participants' personal information was collected with questions about age, gender, nationality, and education. It took participants about 5 minutes to complete the survey.

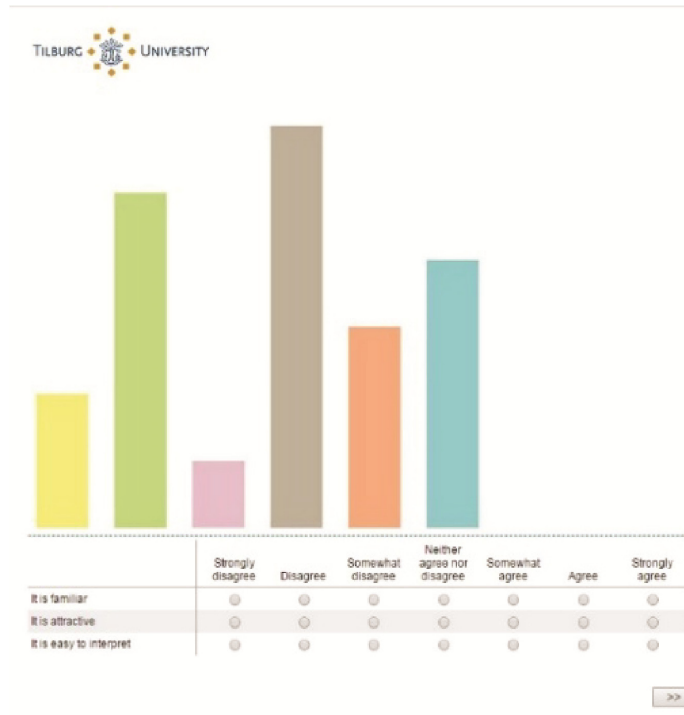


Figure 3 Screen example from survey

4.2.2 Results

Table 1 shows the ratings for each of the three variables for laypeople and designers, ranked from the highest to the lowest ratings. The results confirmed our hypotheses.

Familiarity and perceived ease of use ratings

As becomes apparent from Table 1, the graph types that were used most frequently in our mass media survey are highest in the familiarity rankings, followed by the less frequently encountered types.

- 1 The designers' familiarity ratings were overall higher than the laymen's ratings (designers: $M = 4.99$, $SD = .80$; laypeople: $M = 4.61$, $SD = 1.04$; $t(139) = 2.13$, $p = .021$).
- 2 Designers and laypeople judged the familiarity of the graphs similarly, as evidenced by strong correlations between the ranking orders according to their familiarity ratings: $r = .95$, $p < .001$. Designers and laypeople judged the ease of use of the graphs similarly as well, as the correlation between their ranking orders shows: $r = .95$, $p < .001$.
- 3 Familiarity and perceived ease of use appeared to be positively correlated, both for designers and for laypeople (designers: $r = .55$, $p < .001$; laypeople: $r = .65$, $p < .001$).

Familiarity		Ease of interpretation				Attractiveness	
Laypeople	Designers	Laypeople	Designers	Laypeople	Designers	Laypeople	Designers
6.52 (.82)	6.86 (.35)	6.53 (.69)	6.57 (.85)	5.71 (1.22)	5.39 (1.57)		
6.44 (1.02)	6.84 (.43)	6.34 (1.03)	6.09 (1.18)	5.66 (1.27)	5.09 (1.25)		
5.27 (1.52)	5.80 (1.49)	5.62 (1.29)	5.45 (1.37)	5.12 (1.38)	5.02 (1.47)		
5.09 (1.58)	5.66 (1.48)	5.37 (1.21)	5.32 (1.54)	5.08 (1.28)	4.86 (1.96)		
4.85 (1.64)	5.25 (1.56)	5.13 (1.51)	5.11 (1.62)	4.81 (1.48)	4.75 (1.79)		
4.35 (1.67)	5.00 (1.64)	5.01 (1.55)	5.09 (1.46)	4.81 (1.51)	4.70 (1.81)		
4.12 (1.85)	4.68 (2.00)	4.85 (1.69)	4.48 (1.84)	4.78 (1.47)	4.68 (1.46)		
4.09 (1.90)	4.55 (1.70)	4.75 (1.75)	4.39 (1.65)	4.63 (1.62)	4.64 (1.94)		
3.84 (2.00)	4.41 (1.88)	4.43 (1.72)	4.30 (1.84)	4.56 (1.78)	4.48 (1.80)		
3.79 (1.93)	4.20 (1.65)	4.27 (1.67)	3.86 (1.86)	4.38 (1.85)	4.48 (1.53)		
3.71 (1.78)	3.34 (1.60)	3.62 (1.79)	3.20 (1.86)	3.97 (1.83)	4.45 (1.80)		
3.28 (1.84)	3.30 (1.88)	3.55 (2.05)	3.09 (1.80)	3.93 (1.73)	3.91 (1.70)		

Table 1 Mean attractiveness, familiarity, and ease of interpretation ratings per graph for laypeople and designers (std.dev.) on a scale of 1 – 7

4a *Influence of familiarity and ease of use on attractiveness.* The ranking orders in Table 1 suggest that designers and laypeople differ in their preferences for certain graph types. In order to investigate the relationships between familiarity and perceived ease of use, and attractiveness, we performed correlational and regression analyses.

For both the laypeople and the designers, the mean familiarity and attractiveness ratings were positively correlated (laypeople: $r = .58, p < .001$; designers: $r = .26, p < .001$), as were the mean ease of use and attractiveness ratings (laypeople: $r = .47, p < .001$; designers: $r = .56, p < .001$). Because of the correlation between familiarity and perceived ease of use, we also analyzed partial correlations to see how both factors contributed to attractiveness separately. For the laypeople we found positive correlations again, both between familiarity and attractiveness ratings ($r = .28, p < .001$) and between ease of use and attractiveness ratings ($r = .44, p < .001$). For the designers we found positive correlations between ease of use and attractiveness ratings ($r = .33, p < .001$), but no statistically significant correlations between familiarity and attractiveness ($r = .05, p = .263$)

In order to assess for both groups how much variance in the ratings of attractiveness can be explained by each of the predictor variables, we performed a regression analysis, separately for laypeople and designers. Results are shown in Table 2.

	Laypeople			Designers		
	B	SE B	Beta	B	SE B	Beta
Constant	1.58	0.11		2.84	0.21	
Familiarity	0.24	0.03	.28*	0.24	0.04	.05
Perceived ease of use	0.43	0.03	.47*	0.34	0.04	.38*
	R ² = .46. *p < .001			R ² = .17. *p < .001		

Table 2 Summary of multiple regression analyses for variables predicting attractiveness for laypeople (N=97) and designers (N=44)

Familiarity and ease of use explain a significant amount of the variance in the attractiveness values for both groups, but stronger for laypeople than for designers. The analyses show that for laypeople both familiarity and ease of use significantly explain attractiveness. For the designers perceived ease of use explains attractiveness, but familiarity did not reach statistical significance.

4b Relationship familiarity and attractiveness. In order to test if the relationship between familiarity and attractiveness is linear or curved, we first analyzed the scatterplots of the relationships for both groups (Figure 4). Whereas laypeople’s ratings are best fitted with a linear model, the designers’ ratings are most adequately described by a quadratic model.

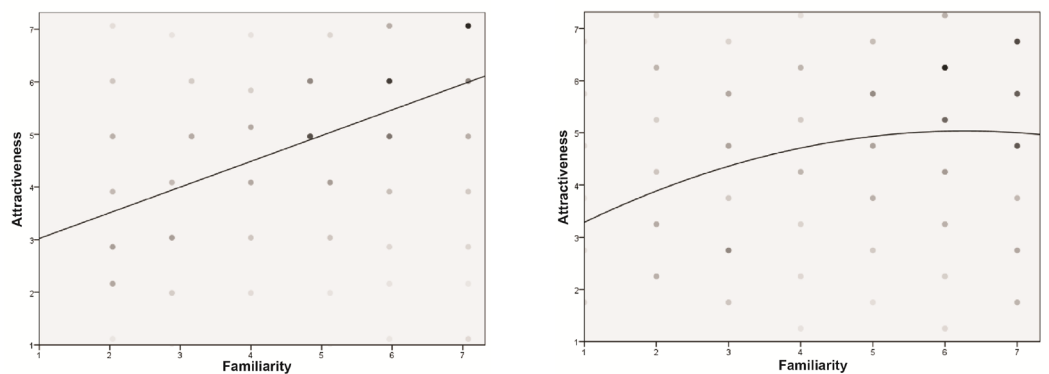


Figure 4 Scatterplots of the relationship between familiarity and attractiveness ratings for laypeople (left) and designers (right)

Curve estimation confirmed that the relationship between familiarity and attractiveness is linear for the laypeople: adding a quadratic term did not cause a significant change in the model fit for them ($t(1161) = .72, p = .471$). For the

designers however, adding a quadratic term did cause a significant change in the model fit ($t(525) = -2.63, p = .009$). A negative beta weight of $-.63$ also suggested an inverted U-shape in the relationship curve. Therefore, we performed a hierarchical regression analysis with familiarity ratings as first predictor and quadratic familiarity ratings as second predictor variable of attractiveness. Adding the quadratic function caused a significant R^2 change ($.01, p = .009$), as table 3 shows. With the linear model alone familiarity predicts 6.8% of the attractiveness values ($R^2 = .068$), whereas with the combination of the linear and quadratic model familiarity predicts 8% of the attractiveness values ($R^2 = .081$).

This result suggests that designers prefer designs situated between the most familiar and the most novel, as we expected.

		B	SE B	Beta
Model 1: linear	Constant	3.53	0.20	
	Familiarity	0.24	0.04	.26**
Model 2: linear and quadratic	Constant	2.56	0.42	
	Familiarity	0.79	0.21	.88**
	Quadratic familiarity	-0.06	0.02	-.63*

$R^2 = .07$ for model 1, R^2 change = $.01$ for model 2 ($P = .009$). * $p < .05$, ** $p < .001$

Table 3 Comparison of linear and quadratic regression models for familiarity predicting attractiveness for designers (N=44)

4.2.3 Discussion

The results of study 1 show that designers and laypeople share ideas about the familiarity and ease of use of the different graph designs, but differ in their appreciations. Familiarity and ease of use are predictors of attractiveness, but differently for laypeople and designers (see Figure 5).

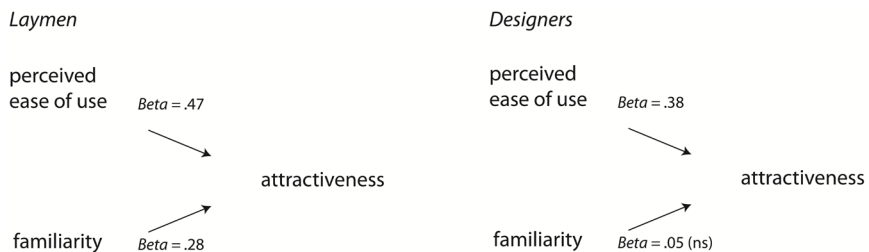


Figure 5 Relative weights of familiarity and perceived ease of use in the prediction of attractiveness for laypeople and designers

Laypeople are attracted to designs that they find familiar and easy to use. Familiarity and ease of use together account for nearly half of the variance in attractiveness ratings. For designers familiarity plays only a minor role in

judgments of attractiveness. Perceived ease of use plays a more important role, but still, familiarity and ease of interpretation together only account for 17% of the variance in the designers' attractiveness ratings. Obviously, we should look for other factors to explain designers' preferences.

Further, the results show that the relationship between familiarity and attractiveness is linear for the laypeople, supporting processing fluency theories that predict that people are attracted to familiar things, and contradicting classical aesthetic theories that predict that people are attracted to an equilibrium between familiarity and novelty. In the case of the designers the relationship between familiarity and attractiveness is quadratic, which suggests that they are attracted to moderately familiar designs. It could be argued that this supports the balance-between-extremes theories. However, these theories do not assume differences between experts and novices, and predict that novices (laypeople) are attracted to moderately familiar designs as well. Therefore, we think it is more plausible that expertise is a moderating factor in the processing fluency theory, as Reber et al. (2004) also suggest. As also shown in other studies (McWhinnie, 1968), experts are apparently attracted to more novel stimuli, despite the pleasure of processing fluency that familiar stimuli seem to offer.

From study 1, all in all, the results show positive relationships between perceived ease of use and attractiveness, although much stronger for the laypeople than for the designers. We wondered how both groups would judge the graph designs after having used them, which we tested in study 2.

4.3 Evaluation study 2

The goal of study 2 was to find out if attractiveness ratings would be influenced by actual use of the graphs in an information retrieval task. We designed a task that reflects the way people commonly use this type of graphs in mass media, namely comparing magnitudes. We asked participants to carry out comparison tasks with the graphs, and then to rate each graph's attractiveness. We did not ask participants to judge the attractiveness before as well as after using the graphs, since asking this twice could have given them too much information about the goal of the study and could therefore have biased their responses.

4.3.1 Method

Participants

Participants in study 2 had similar educational backgrounds and ages as in study 1. Participants were 30 laypeople in design (bachelor and premaster students communication and information sciences at Tilburg University, receiving credits for participation, 17 female, 13 male, mean age 22), and 31 volunteering design specialists (26 graduating bachelor students majoring in graphic design at the art academy of Avans University of Applied Sciences and

4 professional graphic designers, 17 female, 13 male, mean age 24). None of the participants had taken part in the first survey.

Task

The information retrieval task involved comparing magnitudes of components. This task was chosen to reflect the way people normally use these kinds of graphs in mass media (think again of graphs showing election results). The task consisted of a series of statements with quantitative information paired with graphs. Participants were asked to assess if a graph correctly represented the statement (yes or no). We designed statements requiring two types of tasks. The first required a direct comparison between one component of the graph and two others (e.g. 'Fewer sheep have been exported than goats, and more sheep than cows.'). The other type required an indirect comparison: mentally combining two components and comparing the sum with another component (e.g. 'More goats have been exported than sheep and cows together.').

Materials

The same 12 graphs were used as in study 1. In the information retrieval task each graph was accompanied with a legend in the top right corner with the names of the components and the corresponding colors of the components in left-right or clockwise order. For each graph the same color palette was used. No axes or numbers were included, to make sure that differences in performance and appreciation would only be attributed to differences in visual structure, and not to differences in the way axes and numbers were integrated in the graphs. Each graph contained six segments (categories), each representing a different magnitude. The magnitudes were 5, 10, 15, 20, 25, and 30. Both the order of the categories (i.e., the magnitudes) and the colors of the segments were varied to prevent predictability and learning effects. Four different orders were used, evenly spread over graph types and task types. Graphs including legends measured 810 (width) x 500 (height) pixels and were displayed on a 16 inch 1366 x 768 resolution LCD screen together with the statement and yes/no buttons below them. (See Figure 6 for an example.) Like in study 1, attractiveness was measured by a rating task, using again a 7-point Likert scale.

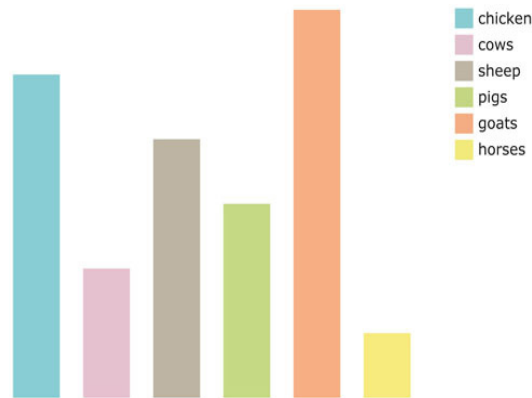


Figure 6 Example of graph with legend

Procedure

Each participant received 24 statements, one direct and one indirect comparison statement for each graph type, with true and false statements evenly distributed among graph and statement type. Graph-and-statement pairs were presented in random order. The information retrieval task started with a written instruction explaining that the participant was about to see a series of statements paired with graphs, one pair at a time, and that it was his/her job to decide if the graph correctly represented the information in the statement or not. Participants were instructed to answer as accurately and quickly as possible. The instruction was followed by two trials. Each statement was presented together with the representing graph until the participant pressed the yes or no button, without time constraints. Statements and graphs were presented in random order. After the information retrieval task, respondents were asked to rate the attractiveness of each graph. This second task also started with a written instruction, explaining that the participant would be presented with the graphs again, one at a time, and that they had to indicate how attractive they found it. Graphs were presented in random order again, one by one until the participant had rated it and pressed the forward button. It took participants about ten to fifteen minutes to complete the test. Usability was measured by logging correct response rates, response times, and response times for incorrect answers. The response times for erroneous responses were used as an indication of the amount of time it takes a participant to complete a seemingly difficult task.

4.3.2

Results

Average correct response rates per graph ranged from 100 to 77 percent for direct, and 93 to 52 percent for indirect comparisons. Average response times per graph ranged from 12 to 27.4 seconds for direct, and 11.4 to 24.8 seconds for indirect comparisons. As we were interested in the question of how the experience of using the graphs would affect attractiveness ratings, in the remainder, we only focus on the attractiveness ratings.

We found no correlations between performance measures (correct response rates, response times, and erroneous response times) and attractiveness ratings, for either of the two groups, neither for the direct comparison task, nor for the indirect one.

Table 4 shows the attractiveness ratings for laypeople and designers, including the rankings of the graphs according to attractiveness, both in study 1 (without use) and study 2 (after use).

















































Laypeople			Designers		
Without use		After use	Without use		After use
 5.71 (1.22)		 6.30 (1.95)	 5.39 (1.57)		 5.10 (1.64)
 5.66 (1.27)		 5.87 (1.20)	 5.09 (1.25)		 4.42 (1.69)
 5.12 (1.38)		 4.93 (1.48)	 5.02 (1.47)		 4.26 (1.44)
 5.08 (1.28)		 4.37 (1.38)	 4.86 (1.96)		 4.19 (1.38)
 4.81 (1.48)		 4.33 (1.21)	 4.75 (1.79)		 4.16 (1.73)
 4.81 (1.51)		 4.33 (1.35)	 4.70 (1.81)		 4.16 (1.86)
 4.78 (1.47)		 4.33 (1.27)	 4.68 (1.46)		 3.94 (1.32)
 4.63 (1.62)		 3.73 (1.70)	 4.64 (1.94)		 3.94 (1.69)
 4.56 (1.78)		 3.60 (1.48)	 4.48 (1.80)		 3.77 (1.52)
 4.38 (1.85)		 3.43 (1.61)	 4.48 (1.53)		 3.74 (1.57)
 3.97 (1.83)		 2.57 (1.48)	 4.45 (1.80)		 3.06 (1.57)
 3.93 (1.73)		 2.03 (1.27)	 3.91 (1.70)		 2.39 (1.63)

Table 4 Attractiveness rankings and ratings without and after use, for laypeople and designers (std.dev.)

T-tests revealed that the attractiveness ratings of the participants from study 2 were significantly lower than those obtained from the participants in study 1 (laypeople: $t(359) = -6.75, p < .001$; designers: $t(371) = -8.77, p < .001$), suggesting a moderating effect of actual use on attractiveness.

To further analyze effects of actual use on attractiveness, we calculated the ranking orders of the graphs according to attractiveness ratings and analyzed, for designers and laypeople separately, the correlation between the rankings in study 1 (no use) and study 2 (use). A strong positive correlation between the attractiveness ranking orders would indicate that use had no effect, whereas a weak or negative correlation would indicate that use did

have an effect on attractiveness. The laypeople's rankings with and without use appeared to be significantly correlated ($r = .825$, $p = .001$), suggesting no effect of actual use. Further, the laypeople's attractiveness ranking order in study 2 appeared to be significantly correlated with the familiarity and ease of interpretation ranking orders in study 1 (ease: $r = .782$, $p = .003$; familiarity: $r = .754$, $p = .005$). For the designers we found no correlation between attractiveness rankings in study 1 and 2 ($r = .225$, $p = .482$), which suggests that the attractiveness ranking order in study 2 differed substantially from study 1. Further, the designers' attractiveness ranking order in study 2 showed no correlation with the familiarity ranking orders found in study 1, but it did with the perceived ease of use ranking order in study 1 ($r = .593$, $p = .042$).

4.3.3 Discussion

No relationships were found between attractiveness ratings and performance measures. Participants did not receive feedback about correctness of their responses, so we did not expect relationships between correct response rates and attractiveness. But they could have experienced that with some graphs it took longer to carry out the tasks than with others. Apparently, efficiency, in terms of response times or erroneous response times, did not influence judgments of attractiveness of either of the two groups. The finding that attractiveness ratings in study 2 were lower than in study 1 may indicate a moderating effect of actual use on attractiveness. Further, after having experienced using the graphs, the designers find other graph types attractive than without user experience. The correlation between the attractiveness and perceived ease of use ranking orders suggests that after having used the graphs, not actual, but perceived ease of use influences their judgments of attractiveness.

4.4 General discussion and conclusions

We were interested in the influence of familiarity and perceived ease of use on attractiveness of graphs for designers and laypeople in design.

Familiarity and perceived ease of use appeared to be correlated for both groups. Both variables also appeared to be predictors of attractiveness for both groups, but differently for laypeople than for designers. Both variables accounted for almost half of the variance in the attractiveness ratings of the laypeople, and for less than 20% of the variance in the attractiveness ratings of the designers. Therefore, other factors must be of influence on attractiveness judgments as well. In the field of HCI interesting studies have been done attempting to reveal such factors (Lavie & Tractinsky, 2004; Tuch, Roth, Hornbaek, Opwis, & Bargas-Avila, 2012; Hassenzahl, 2004). Certainly, some of the factors as defined by them will also apply in graph design such as being clear, original, or creative, but others, such as service quality factors (e.g. 'feel

joyful', 'can count on site') may not be very relevant. An interesting direction for future research is to investigate criteria for graph design for a broader public.

Our finding that the relationship between familiarity and attractiveness is linear in the case of laypeople contradicts the classic theories that predict that people prefer moderately familiar stimuli (Hekkert & Leder, 2008; Berlyne, 1971) and supports processing fluency theories that predict that people prefer familiar and thus 'perceived as easy to process' stimuli (Reber et al., 2004). This result is not in line with findings of for example Hekkert et al. (2003), who found that people prefer moderately typical designs (novel, but still typical). An explanation for this may perhaps be found in the type of stimuli. Empirical studies within experimental psychology attempting to find evidence for the equilibrium between extremes theory have shown inconclusive results, whereas studies in product design showed that people prefer unusual yet typical products. In the first type of study, simple artificial stimuli were used. In the latter, daily products such as telephones and teapots were used. Just as in the product design studies, we used realistic stimuli, made for use in everyday life. The difference between the results of Hekkert et al. (2003) and ours may perhaps be caused by differences in the types of stimuli. Consumer products like teapots are quite ordinary and using them does not require much cognitive effort. Graphs on the other hand are less ordinary to most people, and may be perceived as more difficult to process. Even highly familiar graphs might still offer enough arousal to be aesthetically pleasing. A study investigating how novices construct graphs also showed that novices prefer familiar graph types, because, as participants explained, they understood them well (Grammel, Tory, & Storey, 2010). Some studies have indeed shown that the mere exposure-effect is more easily found for complex stimuli than for simple ones (Bornstein, Kale, & Cornell, 1990). This would mean that repeated exposure and thus increasing processing fluency is appreciated for complex stimuli, whereas a balance between familiarity and novelty is appreciated for simple stimuli, that otherwise might become boring after repeated exposure.

The finding that designers prefer more novel designs than laypeople is as expected, and may be explained by their tendency to consider aesthetic value, besides perceived ease of use. In a previous study we asked laypeople why they preferred standard designs and designers why they preferred non-standard designs (Quispel & Maes, 2014). It appeared that laypeople appreciate clarity the most, whereas the designers mentioned attractiveness and being different from the standard as the main reason for their preference. A more elaborate, qualitative study, in which participants are interviewed in more depth about their evaluations of particular types of designs could shed more light on their reasons for aesthetic preferences.

Study 2 revealed that there were no correlations between attractiveness ratings and performance measures for either of the two groups. Actual ease of use apparently did not affect attractiveness in the task we used. Studies in HCI have shown that poor usability may affect evaluations of aesthetics

(Tuch et al., 2012). Perhaps stronger usability manipulations would also affect attractiveness in data visualization.

Of course there are limitations to this study. We found strong correlations between familiarity and perceived ease of use, but judgments of ease of use and attractiveness may also be based on other factors. The tested graphs differ not only in terms of familiarity, but also in terms of the features they use to represent quantity, such as length or area, and in the way the segments are arranged, as separate parts or as parts of a whole. This may also affect perceived and actual ease of use (Cleveland & McGill, 1984; Hollands & Spence, 1998). Besides design characteristics, also reader characteristics and type of task may influence evaluations. All participants had received higher education, but people's understanding is also influenced by culture and experience (Norman, 2004). As for the type of task, we deliberately chose a type of task reflecting daily life kind of graph use, but perceived and actual ease of use could be different with another type of task. Further, both aesthetics and usability are complex constructs, while we only measured correlations between single items. It would be worthwhile to further investigate the relationship between aesthetics and usability by using multi-level scales. As mentioned above, this would require the development of measurement scales that are appropriate for graph design, and that measure criteria considered relevant by both designers and laypeople in design.

Still, our findings show clear differences in appreciations between designers and laypeople, the designers' audience. This means that designers should be well aware of possible differences between their and their audiences' ideas about attractiveness and understandability, or clarity. Further, the fact that the two groups share opinions about perceived ease of use, but differ in their judgments of attractiveness, may also indicate that attractiveness does not affect perceived ease of use, as is suggested in several studies, but that perceived ease of use may affect attractiveness. After all, if attractiveness would predict perceived ease of use, laypeople and designers should not only differ in their attractiveness judgments, but also in their judgments of ease of use, which was not the case. Moreover, the shift in the designers' appreciations after use suggested an influence of perceived, rather than actual usability on attractiveness. A future study should shed more light on the precise relationships between aesthetics and usability in graph design.

Reading graphs

The role of length and area in comparing quantities

Abstract

Studies investigating the usability of bar vs. pie graphs show contrasting results and researchers disagree about which perceptual features are primarily responsible for their effect. In two studies, we offer evidence for the role and the effect of two crucial perceptual features in reading graphs: length and area. We made use of a large collection of graph designs representing nominal and quantitative information, which all crucially depend on the use of length or area to represent quantity. In an evaluation study, we examined which features are actually perceived by non-expert users as representing quantity. Results show that their judgments are less clear cut than the assumptions of researchers. For most graphs, more than one feature was perceived to play a considerable role, and overall, area was perceived to play a more crucial role than length in the majority of graphs. The study provided us with a classification of graph types based on the role of these features. In a second study, we asked respondents to make simple and more complex comparisons between quantities in a graph. Performance was more accurate and efficient with length than with area representing quantity, but only in complex comparisons.

This chapter is based on:

Quispel, A., Maes, A., and Schilperoord, J. (submitted). Reading Graphs. The role of length and area in comparing quantities.

5.1 Introduction

Studies investigating the usability of bar vs. pie graphs show contrasting results. Many scholars and designers warn against the use of pie graphs. In a review, McDonald-Ross (1977) advocated the use of the bar instead of the pie chart, saying that ‘we still know too little about pie charts to feel entirely comfortable at their widespread use’ (p.375). More recently, Tufte (2001) proclaimed that ‘pie charts should never be used’ (p.178), and Few (2004) recommended to avoid the use of pie graphs because in his view graphs that use area to represent quantity (like pie graphs) communicate poorly (p.61). Despite these opinions, empirical studies into the usability of pie versus bar graphs have shown inconclusive results so far. Researchers also disagree about which perceptual features are primarily responsible for the effect of pie and bar graphs. For bar graphs either length or position along a common scale is considered most crucial, for pie graphs, area or angle compete for the status of most crucial feature.

With this study we contributed to this discussion in different ways. First, we took the discussion of perceptual features and graph usability out of the context in which it has been studied thus far, i.e., the realm of expert users, mostly statisticians and scientists. Instead, we focused on everyday graphs used by everyday, non-expert users to carry out everyday information tasks. Second, we did not restrict ourselves to pie and bar graphs only, but studied a larger collection of graph designs exploiting features such as length and area in different compositions. This way, we mimicked the growing variation of graph designs used in mass media nowadays. Third, we collected evidence about the role of perceptual features in graphs which goes beyond mere theoretical assumption or researcher’s intuition. Finally, we investigated the usability of these perceptual features using tasks which match the everyday non-expert use of graphs in mass media.

In the remainder of this section, we first introduce the discussion on the role of perceptual features in reading quantities in graphs, and on the effect of graph design and perceptual features on performing graph reading tasks. Next, we derive from this the goals of the studies presented in subsequent sections.

5.1.1 Comparing quantities in graphs: assessing the role of perceptual features

There are a number of perceptual features, or graphical codes, which can be used to visualize quantities in a graph (e.g. Bertin, 1983). These features represent the ways in which the two dimensions of the ‘plane’ can be employed to encode quantity. Quantities in a graph can be represented by the position of the top of a bar along a common scale, by the length of a bar, by the area of a segment of a pie or by the angle of the slice, by the volume of a graphical object, et cetera. Several researchers have made claims about the relative effectiveness of these features, thereby focusing on features of

well-known graph designs like bar, divided bar, and pie graphs (e.g. Cleveland & McGill, 1984) These claims are based on the researchers' assumptions about which features encode quantity in these graphs, assumptions which not always converge. As for bar graphs, some researchers consider position (of the top of the bars) along a common scale as being most crucial in representing quantity (e.g. Cleveland & McGill, 1984), whereas for others the length of the bars encode quantity most crucially (e.g. Spence & Lewandowsky, 1991). As for pie graphs, some conjecture that the angles of the segments are most crucial (e.g. Cleveland & McGill, 1984; Simkin & Hastie, 1987), whereas for others area is most important (e.g. MacDonald-Ross, 1977; Spence & Lewandowsky, 1991; Few, 2004). Obviously, all these features are relevant in representing quantities in these graphs: changing quantities in a bar graph automatically means changing the length of bars as well as the top positions of the bars. Changing quantities in a pie graph automatically means changing the angles as well as the areas of the proportions. Moreover, changing quantities in bar graphs also results in changing the area of bars, while changing pie graph quantities also results in changing arc lengths.

One of the most influential studies into the effectiveness of perceptual features in graph comprehension is that of Cleveland and McGill (1984). Their starting point is a list of 10 elementary perceptual tasks they claim people perform to extract quantitative information from graphs, e.g. judging position along a common scale, or length, or area. They also hypothetically define which of these tasks are most crucial in different graph designs. Furthermore, they hypothesize a theoretical ordering of these tasks from most to least accurate, based on a mix of 'own reasoning and experimentation with various graph forms, results of psychophysical experiments, and the theory of psychophysics' (p. 537). For the relevant features, they hypothesize the following accuracy ordering: 1. position along a common scale, 2. positions along nonaligned scales, 3. length-direction-angle, 4. area, 5. volume-curvature, 6 shading, color saturation. For their prediction on the accuracy of length vs. area, they refer to a review article of Baird (1970), which shows that people make more accurate estimations of the magnitude of some aspect of a physical object when they are asked to judge length than when they have to judge area. For the ordering of angle and area, no such motivation is given. Starting from these theoretical assumptions, Cleveland and McGill report the set up and results of two experiments, one using bar and divided bar graphs, the other using pie and bar graphs. They asked respondents to compare the quantities of two dotted segments of graphs, and estimate the percentage of the smaller segment compared to the larger segment. In the first experiment, accuracy of the simple (grouped) bar graph was higher than that of the divided bar graph; in the second experiment, the accuracy of the bar graph was higher than that of the pie graph. They take these results to support (part of) their theoretical accuracy ordering of perceptual tasks, starting from assumed connections between graph designs and perceptual features (bar graphs: position along a common scale; divided bar graphs: length; pie graphs: angle).

Experiments like these certainly provide evidence for the relative efficiency of different graph designs, but they are inconclusive as to the role of perceptual features, as the connection between graph designs and graph features (perceptual tasks), though intuitively plausible, remains unmotivated. Researchers disagree about which features are most crucial in different graph designs, and it is unclear to what extent the researchers' assessments are influenced by their experience with graphs.

Starting from the growing popularity of graphs in mass media for the large public, one may ask how non-expert users of graphs look upon the role of different perceptual features in reading graphs. Different methods can be used to determine which features are important in reading off quantity differences in graphs. The most direct way is the use of methods like eye tracking to monitor the behavior of graph readers, as was done by Goldberg & Helfman (2011). They show that eye tracking can provide useful information about the role of particular features in graph reading tasks. At the same time, the study shows that eye tracking is better in discovering reading strategies than in determining the role of specific features. For example, in assessing the importance of length vs. area in reading bar graphs by using eye tracking, it is difficult to define different areas of interest for length and area, as they are part of the same visual plane. A more straightforward method is to present respondents with an evaluation task, in which they are asked to assess the importance of different features in reading graphs. This method does not allow conclusions on the actual use of these features, but offers evaluative information on how important the features are considered to be by the viewers.

The role of perceptual features is not only defined by their intrinsic characteristics (like their designs or conspicuousness), but also by the type of task to be performed (e.g. Carswell, 1992; Hegarty, 2011). Popular tasks in usability research involve estimating the exact proportion in percentages of one segment compared to the whole of segments (e.g. Hollands & Spence, 1998), estimating the exact proportion of one segment compared to one or more other segment (e.g. Simkin & Hastie, 1987), or else to simply compare the relative size of one segment with the size of one or more other segments (e.g. Spence & Lewandowsky, 1991). Estimating exact proportions requires not only perceptual skills, but also the ability to translate this perception into the abstract format of a percentage, while comparing relative magnitudes offers a more direct connection with perception. Also, the estimation of exact percentages does not reflect the use of graphs in daily life, as graphs are usually presented with percentages once these are considered relevant in the given context.

Spence and Lewandowsky (1991) point out that most psychophysical experiments have focused on comparing segments with wholes, and estimating exact percentages, instead of judging relative magnitudes. Experiments using such tasks show contrasting results on the efficiency of the two most standard graph designs used to represent quantities, i.e., bar vs. pie graphs. For example, Simkin and Hastie (1987) asked participants to estimate the percentage of a smaller segment compared to a larger one, and found that bar

charts were more accurate than pie charts, in line with Cleveland and McGill's results for the same task. In a second task, in which respondents had to estimate the percentage of a segment compared to the whole, performance with the bar and pie graphs was equally accurate. In both tasks respondents were fastest with bar graphs. Hollands and Spence (1998) also used a one-to-all exact estimation, but they found that performance with bar graphs was less accurate, and also slower than with pie graphs and divided bar graphs.

As far as these type of tasks is concerned, it is hard to come up with a clear conclusion on the effect of pie vs. bar charts.

5.1.2 Goal of the paper

In this paper, we aimed to collect more evidence on the role and the effect of perceptual features in reading quantities in graphs, in particular in the context of the growing diversity of graphs in mass media aimed at a diverse non-expert audience. Our aim was twofold. In a first study, an evaluation study, we collected evidence on how non-expert users judge the importance of perceptual features in representing quantities in graphs. This provided us with a classification of graph types based on these features. In a second study, we asked respondents to carry out perceptual tasks with these graphs. The two studies enabled us to draw conclusions about the effect of perceptual features on the usability of graphs.

In selecting the graph designs used in the studies, we started from the observation that most of the graphs in mass media represent a combination of nominal and quantitative data, showing different portions of a particular phenomenon, like numbers/percentages of votes for political parties, or asylum seekers in different EU countries (see Quispel, Maes & Schilperoord, 2016). Therefore, in our study we focused on graphs representing a combination of nominal and quantitative data, and excluded designs displaying trends or other ordinal/interval data, which are usually based on line designs (Zacks & Tversky, 1999).

Instead of focusing only on the traditional graph designs (bar and pie graphs), we used a larger collection to reflect the growing diversity of graph designs in mass media. We used more or less familiar designs (bar, pie, divided bar, donut chart, bubble chart, semi-circle) as well as more novel variants. That way, we created a larger variation of ways in which the relevant graphical features are presented and combined, and thus created a more reliable and robust basis to assess the role and effect of these features.

As to the perceptual features we asked respondents to judge in study 1, we did not include all features listed in Cleveland and McGill, but restricted ourselves to the four features most associated with the basic designs of bar and pie graphs respectively: length and position along a common scale on the one hand, and area and angle on the other. The other features mentioned in Cleveland and McGill – position along non-aligned scales, direction, volume,

shading, color saturation and curvature – are not relevant to the graph designs under study.

Finally, in assessing the effect of graph designs and perceptual features in study 2, we asked respondents to carry out tasks which are congruent with the way in which these graphs are used in everyday media: we asked them to compare relative magnitudes. At the same time this type of task fits in with the focus of our study on the role of perceptual features, as this task only solicits the reader's ability to perceptually compare shares, without having to translate them into an abstract format (like an exact estimation of percentages). Spence and Lewandowsky (1991) used similar tasks. They showed participants one graph at a time and asked them to judge if a segment was larger or smaller than another (direct comparison) or than a combination of two other segments (complicated comparison). Their results showed a slightly better accuracy with pie charts but only for the complicated comparisons. The predictive value of these results for our effect study is low, however, as the explanation they give for the beneficial effect of pie charts is not based on area or angle as crucial feature, but on the pie chart's circular shape, which would offers readers imaginary anchors that divide the pie in halves and quarters, an explanation also mentioned in Simkin and Hastie (1987). In our study, we aimed to reduce effects of such other design variables by using a larger variety of graph designs, including variants in which the use of area or angle is not combined with circular shape.

5.2 Study 1: perceptual features evaluation study

In study 1, we conducted a survey to establish which perceptual features are perceived by non-expert graph users as playing a key role in the representation of quantity in a range of different graph designs. We asked respondents in a survey to evaluate the role of four perceptual features in judging quantities when comparing quantities in these graphs. The four features represent the major features in bar or pie like graphs as we used them in the experiment: position along a common scale, length, area, and angle. We used the online survey environment CrowdFlower (www.crowdfunder.com) to administer the survey.

5.2.1 Method

Participants

The online survey was taken by 252 participants in 15 western countries for \$0.30 compensation, ages 15-25 (17.9%), 25-45 (58.7%), 45-65 (22.2%), and 65-100 (1.2%). Their educational levels ranged from none/primary school (1.2%), high school (34.5%), and lower/secondary vocational (16.7%) to higher vocational/university (47.6%).

Materials

As test materials we used the same set of graphs as used in Quispel et al. (2016). Each graph visualized one and the same data set consisting of a combination of nominal and quantitative data (see Figure 1). We used six nominal categories, as it has been argued that this kind of graphs is suitable for up to six categories (Bertin, 1983) and using more categories would be disadvantageous for some graph designs, such as the bar graph (Hollands & Spence, 1998).

The 12 graph designs were chosen to reflect the diversity of different visualization techniques with which a combination of quantitative and nominal data can be represented. The construction of a graph is determined first by the dimensions of the 'plane' (surface) used to represent quantity (e.g. length or area), and second by the particular way in which these dimensions are utilized (rectilinear or circular) (Bertin, 1983). The combination of both results in a variation of possible types of construction, which was represented in the set of graph designs. Six of the graph designs (the upper row in Figure 1) are more or less frequently used in the media (see Quispel et al. 2016, who counted graphs over a period of six weeks in three Dutch newspapers and three magazines, and found 57% bar graphs, 16% pie graphs, 10% divided bars, 8% donut charts, 7% bubble charts, and 2% semi-circles). The other six graph designs (the lower row in Figure 1) are not or hardly ever found in mass media.

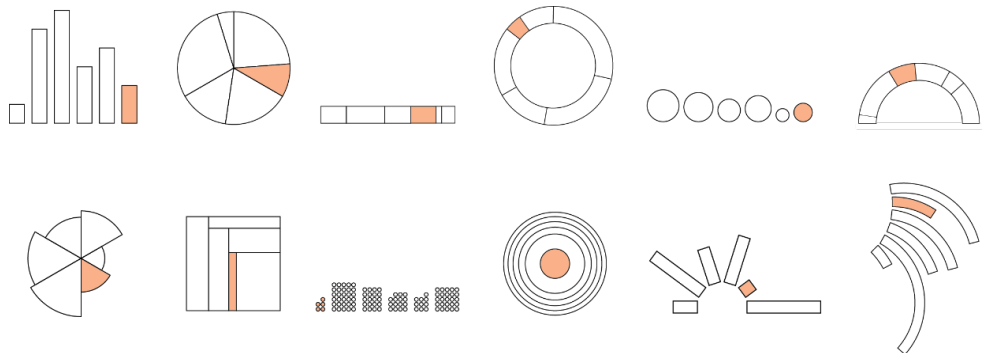


Figure 1 Test materials used in study 1

For each graph, two variants were constructed (see Figure 2), in order to present respondents with 12 pairs of the same graph design. In each pair, the same segment was colored red, but the size of the segment was different. The rest of the graph was designed in black and white.

Each pair of graphs was accompanied by the same four sliders. Each slider measured the importance of one perceptual feature (position along a common scale, length, area, angle) in the graph design presented above. See Figure 2. Sliders had a scale of 0 to 100, with zero meaning that a feature played no role, 100 that the feature was perceived as extremely important in comparing quantities in this graph. So, participants could score each feature on a 0-100 scale.

Compare the quantities of the red categories. Indicate for each feature how important it is in the comparison.

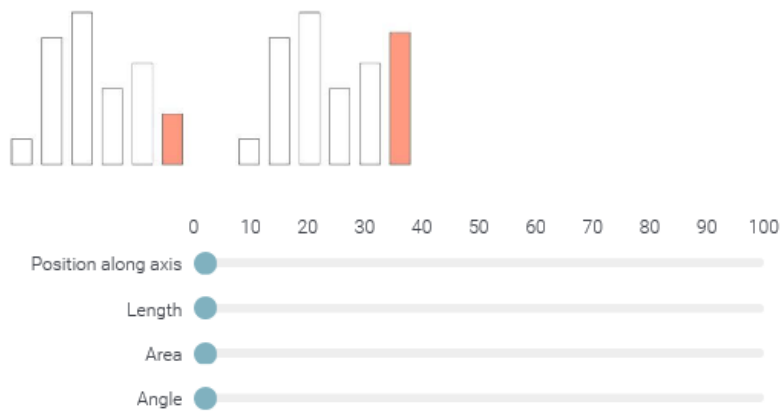


Figure 2 Example screen from study 1 consisting of a pair of bar graphs, together with the four evaluation questions

Procedure

The survey started with an instruction explaining that in certain graphs quantities can be represented by position along a common scale, area, angle, and length, each possibility illustrated by an example (different from the tested graphs). Then participants were explained that they were about to be presented with a series of graph pairs, which were identical except for one red colored segment, as illustrated in Figure 2. Participants were asked to compare the quantities of the red categories in the two graphs, and then to indicate how important they considered each of the perceptual features – length, area, position, angle – by using the four sliders. After having rated the importance of the four features, respondents had to press the forward button, after which a new graph pair appeared.

Participants were instructed that there are no correct or incorrect answers, and that therefore no feedback would be given. The survey could only be completed when each question was answered. The survey took about 5 minutes to complete. After this assessment task, participants' personal information was collected with questions about age and education.

5.2.2

Results

Table 1 shows to what extent each of the four features is perceived as playing a role in judging quantity on a scale of 1 – 100. A first observation is that in all graphs more than one feature is to some extent perceived as relevant in representing quantity. A second observation is that in each graph either length or area received the highest score. Finally, in most graphs (n=8) area is considered most important.

The comparison of the mean scores for the four features in each graph shows that in 3 cases, including the bar graph, the scores for length, and in

7 cases, including the pie graph, the scores for area were significantly higher than the scores for the other three features ($p < .001$); these scores are marked bold in Table 1. In the remaining two cases the scores for length and area were significantly higher than for the other two features ($p < .001$), but they did not differ significantly from each other. Based on these results, we distinguish three groups of graphs, as can be seen in the ordering of Table 1: graphs perceived to rely on length ($n=3$), on area ($n=7$) or combined ($n=2$).

Even if one feature is perceived as significantly more important than the others, the graphs differ in the degree to which more than one feature is considered important. For example, in the case of the bubble chart, the score for area is relatively high compared to the scores for the other three features. On the other hand, in the case of the pie graph, scores are high for area, but angle is perceived as playing a relatively large role as well. In other words, some graphs are perceived to rely more on one single feature, while others rely more on two features. In order to be able to control for this difference, we calculated for each graph the proportion of the highest feature score in the total of scores for the four features. At the same time, high standard deviations indicate that respondents are not very homogeneous in their judgments. To take this variation in judgments into account, we first recalculated the scores for each feature by dividing them by their standard deviations. The scores thus obtained were used to calculate the proportion of the highest value of the total of scores, as shown in Table 1. The scores give us a degree of 'singularity': the higher the score, the more a graph is judged to be relying on a single feature, the lower the score, the more it is dependent on a hybrid of features. The results do not give us reason to change the three classes distinguished above, but they certainly support the idea that graphs in general are judged to be dependent on hybrids of features.

Finally, participants in the survey had varying levels of educational background and therefore probably varying degrees of experience with reading graphs. To see if graph reading experience might have influenced the results, we separately analyzed the results for participants with higher education (higher vocational/university; $N = 120$). Results showed that exactly the same graphs were perceived as using length and area to represent quantity, and in the same two graphs length and area were considered to be about as important in the comparison of magnitudes.

5.2.3

Discussion

In 10 of the 12 graph designs either length or area is perceived as the key perceptual feature in comparing magnitudes, with length being most important in the bar graph and area in the pie graph. This shows an interesting difference between opinions of expert and non-expert users. Apparently, for non-expert users, the features length and area are judged to be more crucial than position along a common scale and angle. In most of these cases ($n=7$) area is perceived as the principal feature. The preference for area and length over angle

Reading graphs

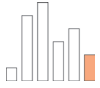
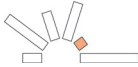










	position	length	area	angle	shares/std. dev.	singularity
	Share per feature (std. dev.) share divided by std. dev.					
	36.41 (34.4) 1.06	71.38 (31.0) 2.30	39.98 (21.2) 1.89	21.18 (31.5) 0.67	5.92	39
	29.29 (32.5) 0.90	67.61 (29.4) 2.30	49.35 (33.5) 1.47	30.43 (33.4) 0.91	5.58	41
	26.85 (31.9) 0.84	60.65 (31.3) 2.30	44.31 (33.5) 1.32	34.19 (32.9) 1.04	5.50	42
	20.65 (29.5) 0.70	22.14 (28.2) 0.79	72.65 (30.0) 2.42	19.35 (29.3) 0.66	4.35	56
	24.40 (28.5) 0.86	33.18 (32.8) 1.01	72.28 (29.6) 2.44	20.62 (29.4) 0.70	5.01	49
	30.79 (35.3) 0.87	20.25 (27.5) 0.74	71.08 (30.1) 2.36	17.56 (28.1) 0.62	4.59	51
	27.04 (30.9) 0.88	26.60 (28.4) 0.94	68.76 (30.9) 2.23	50.81 (35.4) 1.44	5.49	41
	24.50 (30.8) 0.80	31.81 (31.3) 1.02	68.73 (36.8) 1.87	41.53 (36.8) 1.13	4.82	39
	28.57 (32.5) 0.88	33.91 (30.7) 1.10	65.54 (31.8) 2.06	18.91 (26.7) 0.71	4.75	43
	26.02 (31.0) 0.84	49.63 (34.8) 1.43	60.34 (32.6) 1.85	33.33 (34.1) 0.98	5.10	36
	32.98 (32.3) 1.02	63.71 (30.1) 2.12	58.80 (31.2) 1.88	19.77 (29.2) 0.68	5.70	37
	26.96 (31.0) 0.87	53.47 (31.8) 1.68	56.11 (31.3) 1.79	33.97 (32.7) 1.04	5.38	33

Table 1 Perceived share on a scale of 0 – 100 of each graph feature in representing quantity per graph (std.dev.), shares divided by their standard deviations, and singularity values for each graph (proportion of the highest share of the total of shares, divided by standard deviations)

and position along a common scale may be explained by their prima facie conspicuousness. Length and area are basic intrinsic characteristics of all objects and inspire many conceptual processes based on elementary primary metaphors like up is more, and larger is more (e.g. Lakoff & Johnson, 1980; Schubert, 2005). The preference of area over length may partly be explained by a skewed selection of the 12 graphs. But the dominance of area is consistent with the idea that area is determined by two dimensions of the plane and thus more readily perceivable for non-expert viewers.

In two cases length and area are considered equally important. Interestingly, one of these two graphs is the divided bar graph. Apparently, as for bar and pie graphs, also for the divided bar graph non-expert judgments differ from the researchers' opinions, since Cleveland and McGill consider length the crucial feature in the divided bar graph. The role of area is considered as important as length in the semi-circle, which may be considered a circular variant of the divided bar graph. And in the case of the 'donut', also a circular variant of the divided bar graph, the score for area is even higher than for length.

Starting from a classification of graphs in three types (length-based, area-based, hybrid) the question is how these types perform in comparison tasks carried out by non-expert users. This was investigated in study 2.

5.3 Study 2: Experiment²

Our aim in study 2 was to investigate the performance of the three types of graphs: graphs perceived to be based on length (L-graphs), area (A-graphs) or both (LA-graphs). We conducted an experiment with an information retrieval task involving two types of comparisons between segments in a graph.

5.3.1 Method

Participants

Participants were 30 students Communication and Information Sciences at Tilburg University, for credits. (17 female, 13 male, mean age 22). None of the participants had taken part in the graph features or familiarity rating survey.

Task and materials

As test materials we used the 12 graphs that were used in study 1. We distinguished 3 graph feature types: Length based (L, n=3), Area based (A, n=7), Length-Area based (LA, n=2), based on the results of study 1. The information retrieval task involved comparing magnitudes of segments. The task consisted of giving a yes-no answer to a comparison statement paired with a graph. For an example, see Figure 3. Participants were asked to assess if a graph correctly represented the statement (yes or no). We designed comparison statements with two complexity levels. The first required a direct comparison between one segment of the graph and two others (e.g. 'Fewer sheep have been exported than goats, and more sheep than cows.'). The second required an indirect comparison: mentally combining two segments and comparing the

2 This study is a reanalysis of part of study 2 in Quispel, Schilperoord & Maes (2016), in which these tasks were presented to two groups of respondents (experts and laymen in design). Given the setup of the current study, only the laymen results were used, though the results do not differ significantly for the two groups.

sum with another segment (e.g. 'More goats have been exported than sheep and cows together.').

Each graph was accompanied with a legend in the top right corner with the names of the components and the corresponding colors of the components in left-right or clockwise order. For each graph, the same color palette was used. No axes or numbers were included, to make sure that differences in performance and appreciation would only be attributed to differences in visual structure and not to differences in the way axes and numbers were integrated in the graphs. Each graph contained six segments (categories), each representing a different magnitude. The magnitudes were 5, 10, 15, 20, 25, and 30. Both the order of the categories (i.e. the magnitudes) and the colors of the segments were varied to prevent predictability and learning effects. Four different orders were used, evenly spread over graph designs and task types. Above each graph, at the top of the screen the yes-no statement was presented.

Graphs including legends measured 810 (width) x 500 (height) pixels and were displayed on a 16 inch 1366 x 768 resolution LCD screen together with the statement and yes/no buttons below them.

See Figure 3 for an example of a graph with legend and statement as it appeared in the survey.

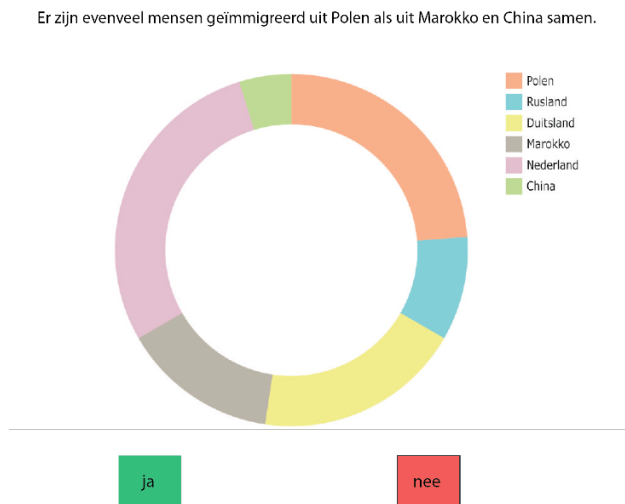


Figure 3 Example of a graph with a legend and statement

Procedure

Each participant received 24 statements, one direct and one indirect comparison statement for each graph design, with true and false statements evenly distributed among graph and statement type. Graph and statement pairs were presented in random order. The information retrieval task started with a written instruction explaining that the participant was about to see a series of statements paired with graphs, one pair at a time, and that it was his or her job to decide whether the graph correctly represented the information in the

statement or not. Participants were instructed to answer as accurately and quickly as possible. The instruction was followed by two example trials. Each statement was presented together with the representing graph until the respondent pressed the yes or no button, without time constraints. After the information retrieval task was completed, respondents were asked to rate the attractiveness of each graph, a task which is not relevant to the current discussion (see Quispel et al., 2016). In all, it took participants about 10–15 min to complete the test. Usability was measured by logging correct response rates and response times.

5.3.2

Results

Table 2 shows correct response rates (%) and response times (ms) per task type (direct, indirect) for the individual graph designs, as well as the average for the three graph feature types.

Accuracy

A 3 x 2 ANOVA with graph feature (L, A, LA) and task type (direct, indirect) as within-subject factors revealed a main effect of task ($F(1, 29) = 2.07, p < .001$) on correctness. Answers were more accurate in the direct task than in the indirect task. Further, we found a main effect of feature ($F(2, 58) = 5.03, p = .010$) on correctness, and a trend of an interaction between feature and task ($F(2, 58) = 3.02, p = .057$). Because of this interaction we performed a split analysis. Results showed no effect of feature in the direct task ($F(2, 58) = 0.15, p = .862$), but in the indirect task we found a main effect of feature on correctness ($F(2, 58) = 5.75, p = .005$): answers with length based graphs were more accurate than with area or length-area.

Efficiency

A 3 x 2 ANOVA with graph feature (L, A, LA) and task type (direct, indirect) as within-subject factors revealed a main effect of task on response time ($F(1, 29) = 2.07, p < .001$). Respondents were faster in the indirect task than in the direct task. Results also showed a main effect of feature on efficiency ($F(2, 58) = 10.48, p < .001$) and a significant interaction between feature and task ($F(2, 58) = 7.72, p = .001$). Because of this interaction we performed a split analysis. Results showed no effect of feature in the direct task ($F(2, 58) = 2.60, p = .083$). In the indirect task there was a main effect of feature on efficiency ($F(2, 58) = 17.73, p < .001$): performance with length and length-area graph types was faster than with area types.













	correct response rate		response time	
	direct	indirect	direct	indirect
Length (L)	.92 (.19)	.82 (.24)	15930 (3971)	12799 (3581)
	.96 (.19)	.80 (.41)	17029 (5977)	11385 (3805)
	.89 (.32)	.93 (.25)	17623 (5743)	13509 (4214)
	.90 (.31)	.72 (.46)	13931 (3850)	13633 (6203)
Area (A)	.90 (.11)	.67 (.16)	15410 (3240)	15700 (3801)
	.86 (.35)	.77 (.43)	19033 (5267)	13780 (2990)
	1.0 (.00)	.50 (.51)	12011 (3062)	16725 (6691)
	.93 (.26)	.90 (.31)	14656 (6060)	12297 (3101)
	.77 (.43)	.55 (.51)	19173 (6553)	17383 (4858)
	.90 (.31)	.63 (.49)	14230 (3373)	17732 (7711)
	.87 (.35)	.64 (.49)	15312 (4828)	18044 (6055)
	.97 (.18)	.69 (.47)	13539 (3751)	14339 (4684)
Length-Area (LA)	.90 (.24)	.58 (.37)	14479 (4179)	12874 (4643)
	.93 (.25)	.52 (.51)	14390 (4129)	12891 (5167)
	.86 (.35)	.67 (.48)	14522 (5411)	12980 (5390)

Table 2 Mean percentage of correct responses (%) and response times (ms) per individual graph design and per graph feature type (L, A, LA) (std.dev.).

Singularity

We conducted a 2 x 2 ANOVA with singularity (low, high) and task type (direct, indirect) as within-subject factors. This way, we aimed to measure the effect of the degree of singularity as explained in Study 1. In this analysis we only included the 7 area graph types, as they showed a large variation on that point, ranging from 36 to 56 (as compared to length based graphs, with values between 39 and 42, and LA graphs, with the scores 33 and 37.) We divided the 7 A-graphs in two groups: low singularity (36-46) and high singularity (46-56), see Table 1. Results showed no effect of singularity, neither on accuracy ($F(1, 29) = .10, p = .756$), nor on efficiency ($F(1, 29) = .79, p = .381$), and no interaction between singularity and task type ($F(1, 29) = .24, p = .632$).

5.3.3 Discussion

Not surprisingly, accuracy was higher in the direct task than in the indirect task. The fact that the direct task took more time than the indirect task is an artifact of the setup of the study: the time needed to read the comparison sentences is included in the efficiency measures. Direct tasks were packed in syntactic constructions which were systematically longer than indirect tasks (direct: 'Fewer sheep have been exported than goats, and more sheep than cows'; indirect: 'More goats have been exported than sheep and cows together').

It appears that when people simply have to estimate if one magnitude is larger or smaller than another (i.e., the direct comparison task), it makes no difference if length or area is used to encode quantity. Answers in the L and A graphs were as accurate and fast. Only in the indirect, and mentally more complex comparison task answers are more accurate with L graphs than with A and LA graphs. Further, for the more complex task, L and LA graphs were faster than A graphs. So, for indirect tasks, LA graphs were equally quick as L graphs, but less accurate.

All in all, the results show that performance with the graphs that are perceived as using area to encode quantity is worse than with the graphs that are perceived as using length to represent quantity, be it only in more complex tasks involving summations of magnitudes.

5.4 General discussion and conclusion

In study 1, we investigated which perceptual features are actually perceived by non-expert users as the principal features representing quantity in a number of different graphs. The results nuance the assumptions made by researchers. First, in judging bar and pie graphs, our survey shows a clear preference for length and area, rather than position along a common scale and angle as most important feature. Furthermore, in most of the graph designs either length or area is perceived as the main feature encoding quantity. In the case of two designs it remained inconclusive which of these features is more important: the divided bar, and the semi-circle. Interestingly, in Cleveland and McGill's (1984) study, claims about judging length being more accurate than judging area are based on experiments in which the divided bar graph is considered to use length to represent quantity, not area. Finally, the singularity results show a considerable variation in the degree in which graph designs are felt to be dependent on one or more features. Another observation is that in most cases – 7 out of 12 – area is perceived as the principal feature encoding quantity, and in two more cases area is perceived as equally important as length. Although one may find other outcomes for other selections of graph designs, there are some designs for which it is striking that area is considered most important. For example, the question is why area is perceived as more important than

length in the donut design. The importance of area in this case may be related to the circular shape. In general, it is notable that all three length types are graphs in which the components are arranged as separate parts. It seems that area is more readily perceived as the most or at least equally important second feature representing quantity when the components of the graph are arranged as parts of a whole, as in the semi-circle, the divided bar, and the donut.

The main conclusion of the usability study is the relevance of task type, i.e., the difference between direct and indirect comparisons, reflecting a different level of complexity. For a simple comparison task, the comparison of one segment with one other segment, which can be carried out at one's own pace, length or area does not seem to be a relevant feature. In the more complex comparison task, length based graphs yielded faster and more accurate responses than using area based graphs. The use of a larger set of graph designs makes us quite confident to conclude that these effects are caused by length rather than position along a common scale. It is unlikely that a common scale is relevant in case the bars or the common scale is bended, as suggested also by the results in Table 1. The indirect task implies a mental summation of segments. For this task, it is likely that length, being based on one dimension, offers viewers an easier template than area, which is based on two dimension. At the same time, this explanation makes clear that the advantage we found for length in this study does not necessarily extend to other complex tasks.

The results in Study 2 are not in line with those of Spence and Lewandowsky (1991), who found that performance with the pie graph (an area type according to our survey results) was at least as good as with the bar graph (a length type according to our survey results) in a similarly complex task. The reason for this, as also argued by Spence and Lewandowsky, may be that the pie graph has specific characteristics besides using area to encode quantity (e.g. being round, offering a center point around which components can be turned around) that offer advantages in summation operations over the bar graph. As we also used area based graphs not having these perceptual benefits, the beneficial effect found for pie graphs in Spence and Lewandowsky is likely due to such other features than to the use of area to represent quantity.

For two graph designs, length and area were considered equally crucial. In the complex task, these designs performed less well than length based graphs. One may reason that it would be beneficial to be able to rely on an additional feature (area) on top of the successful feature of length, but the results suggest that a combination of features is not beneficial.

Despite the large selection of designs, confounding factors can still have played a role. All length based graph designs arrange their components as separate segments (as in a bar graph) while the segments in most of the area based graphs are arranged as parts of a graphical whole (as in a pie graph). According to several theories and studies (Freedman & Shah, 2002; Simkin & Hastie, 1987) the type of task we used, comparing magnitudes, would be advantageous for graphs in which segments are arranged as separate parts.

This arrangement should make comparing segments easier. This could mean that the good performance with length based graphs may partly be explained by the separate parts design. However, our area based graphs also included designs in which the segments were arranged as separate parts. Their performance in the indirect tasks was considerably worse than performance with the length based graphs.

All in all, our studies show that it is worthwhile to investigate which features are actually perceived as representing quantity in graphs, which may perhaps be done more elaborately and with more objective methods such as eye tracking. They also show that performance with those features is largely dependent on the type of task. Clearly, our results only apply in the specific type of task we used, involving comparison of relative magnitudes. It would be interesting to investigate the effect of perceptual features in other cognitive tasks as well, again using a wider variety of graph designs than the commonly used bar, divided bar, and pie graphs.

Visual Ability in Navigation Communication

Abstract

The aim of this study was to test the difference between visual artists and engineers with respect to visual-object and visual-spatial ability, and to investigate if this difference results in different strategies for processing and producing navigational information. Respondents carried out description tasks based on route images and drawing tasks based on route descriptions. Results show that object visualizers (designers) process route information more often from a survey perspective, focus more on objects (landmarks) and create more detailed drawings than spatial visualizers. Conversely, spatial visualizers (engineers) process route information more often from a route perspective, focus more on path entities (streets), and create more schematic drawings than object visualizers.

This chapter is based on:

Quispel, A. and Maes, A. (2013). Visual Ability in Navigation Communication. *Information Design Journal*, 20:2, 136-147.

Quispel, A. and Maes, A. (2012). Visual ability in navigation communication. In: *Proceedings of European Association for Research on Learning and Instruction (EARLI) SIG2 'Comprehension of text and graphics'*, bi-annual meeting, Aug 28-31, Grenoble, FR.

6.1 Introduction

Visual communication materials, like visual instructions, maps or user interfaces, are ubiquitous nowadays. We learn from them, we use them as aids navigating in space, we interact with them searching for information or trying to solve problems. We all know the joy a well-designed graphical display can give, and we all once experienced the frustration a bad design can cause. But what makes a good design?

The design of information graphics has been studied extensively, especially in the field of educational psychology and cognitive science, but also in fields in which the use of graphical displays are crucial, like ergonomics, or geography (see for a recent overview Hegarty, 2011). Gradually, the influence of individual differences on the interaction with visual communication materials received more research attention. Many studies include domain or task familiarity (experts vs. novices) as a predictor of performance, but gradually also more fine-grained individual variables, like cognitive style and visual abilities, are taken into account. More insight in these differences and how they affect the use of visual information may help find design factors that could lead to more effective visualizations, which should result in more custom made visualization solutions for people with different abilities or preferences.

6.1.1 Visual intelligence

Since the beginning of the past century a lot of research has been done into individual differences in intelligence. By 1939 researchers had shown that spatial ability is an intelligence factor distinct from verbal intelligence. As of the middle of the 20th century researchers found evidence for several subcomponents of spatial ability (see for a clear overview Hegarty and Waller, 2006). They seem to agree on two classes of spatial ability. One involves the ability to perform complex mental object transformations, called spatial visualization (McGee, 1979; Lohman, 1979), object manipulation (Kozhevnikov & Hegarty, 2001), object based spatial transformation (Zacks, Mires, Tversky, & Hazeltine, 2002), or spatial transformation (Hegarty & Waller, 2003). The other involves the ability to imagine what an object or scene would look like seen from another viewpoint, called spatial orientation (McGee, 1979; Lohman, 1979; Kozhevnikov & Hegarty, 2001), egocentric perspective transformation (Zacks et al., 2002), or perspective taking (Hegarty & Waller, 2003). Visual intelligence has long been considered to consist only of spatial, or visual-spatial, ability. Studies into the facilitative effects of pictures on learning and problem solving led to a distinction between people in terms of their cognitive styles: verbalizers preferring verbal means for processing information, and visualizers preferring visual means for processing information (e.g. Paivio, 1971; Richardson, 1977; Mayer & Massa, 2003). This verbalizer-visualizer dimension was revised by Kozhevnikov, Hegarty, & Mayer (2002) and Kozhevnikov, Kosslyn, & Shephard (2005), who showed that there are two types of visualizers: those with high

spatial ability, called spatial visualizers, and those with low spatial ability, called object visualizers. Object visualizers rely on 'visual imagery' (representations of visual appearances, like shape, color, size), and generate detailed pictorial images of objects. Spatial visualizers rely on 'spatial imagery' (representations of spatial relations of parts of an object, locations of objects in space and their movements) and are good at generating schematic images representing spatial relations among objects and at imagining spatial transformations. This distinction between visual-spatial and visual-object ability is supported by findings in cognitive neuroscience that object properties and spatial properties are being processed by two different pathways in the brain (Farah, Hammond, Levine, & Calvanio, 1988; Kosslyn & Koenig, 1992; Kosslyn, 1994). Kozhevnikov et al. (2005) also found that object visualizers encode and process images holistically, as single perceptual units, whereas spatial visualizers generate and process images analytically, piece by piece. Verbalizers performed at an intermediate level on imagery tasks.

Visual-object intelligence was introduced as a new, independent component of intelligence by Blazhenkova and Kozhevnikov (2010). Defining visual-object intelligence as reflecting one's ability to process information about visual appearances of objects and their pictorial properties (e.g. shape, color and texture), they demonstrated that this type of intelligence is distinct from visual-spatial intelligence, which reflects one's ability to process information about spatial relations and manipulating objects in space (e.g. mental rotations). This study established visual-object ability as a legitimate dimension of intelligence by providing evidence that it satisfies all requirements of an independent dimension of intelligence: unique ecological validity, capacity to support abstract processing, and unique qualitative and quantitative characteristics, irreducible to spatial and verbal components of intelligence. As Blazhenkova and Kozhevnikov point out, the role of visual-object ability is increasing in a wide range of professions and in everyday life, as the use of rapidly presented visual stimuli, which call for quick, holistic processing, are prevailing in contemporary media.

The finding that visual-object ability is a unique form of intelligence, independent from visual-spatial ability, would put an end to the thought that only beta scientists have good visual abilities, and would mean that designers and visual artists are visually intelligent as well. Therefore, this relatively new distinction deserves to be tested more widely.

6.1.2 Visual ability and profession

Ability tests were originally developed as tools for personnel selection, so it is not surprising that certain abilities are associated with performance in certain jobs. Performance on spatial ability tests has since long been associated with performance in engineering, drafting, design, mechanical drawing, art, mathematics, and physics (Pellegrino, Mumaw, & Shute, 1985; Miller, 1996; Kozhevnikov, Motes, & Hegarty, 2007). This is quite a broad category of both

technical and science, and art and craft related professions, which seems to be due to the fact that for a long time imagery was treated as a single, undifferentiated construct. This single visual-spatial ability was thought to predict performance in a variety of professional fields that require any kind of visual thinking (Blazhenkova & Kozhevnikov, 2010).

The recent distinction that Blazhenkova and Kozhevnikov (2010) and Kozhevnikov et al. (2005) found between visual-object and visual-spatial ability has also led to a clearer distinction in the professions correlated with visual-spatial abilities. According to these researchers visual-object ability relates to specialization in visual art and design, whereas visual-spatial ability is related to specialization in science (physics, biochemistry, engineering, computer science and mathematics). Neither visual ability predicts specialization in humanities and social sciences. Specialists in those fields tend to have better verbal abilities.

Several studies support this further distinction in the professions associated with visual-spatial ability. For example, Lindauer (1983) and Rosenberg (1987) found that professional artists are much dependent on their mental imagery, which was described by the participating artists as consisting of vivid, detailed, and colorful images. Aspinwall, Shaw, & Presmeg (1997) found that imagery supports mathematical functioning, but that highly vivid, uncontrollable imagery is a hindrance in constructing meaning for mathematical concepts. A study by Kozhevnikov, Hegarty, & Mayer (1999) showed that while spatial imagery may promote problem solving success in kinematics, the use of visual imagery presents an obstacle to problem solving in this area. Further, Hegarty and Kozhevnikov (1999) found that the use of schematic spatial representations was associated with success in mathematical problem solving, whereas the use of pictorial representations was negatively associated with it.

Interestingly, information visualizations are often designed either by illustrators and graphic designers (visual artists and designers, associated with object visual ability) or by engineers (associated with spatial visual ability). The studies described above show a clear distinction between the two types of visual abilities and their association with these two groups of professionals. However, the consequences of these differences have hardly been investigated yet. How do they show up in the way these two groups of people perform in tasks that require visualizing information or interpreting visual information?

6.1.3 Visual intelligence in navigation tasks

In this study we aimed to replicate the relation between specialization in art and design and visual-object ability and between specialization in engineering and science and visual-spatial ability. Further, we wanted to investigate how this difference is reflected in performance in specific visual communicative tasks. For these tasks we chose the domain of navigation. Visual-spatial ability

has been associated with performance in navigation and in route and environmental learning in numerous studies (e.g. Moffat, Hampson, & Hatzipantelis, 1998; Allen, Kirasic, & Dobson, 1996; Galea & Kimura, 1993; Hegarty, Montello, Richardson, Ishikawa, & Lovelace, 2006; Kozhevnikov, Motes, Rasch, & Blajenkova, 2006; Aginsky, Harris, Rensink, & Beusmans, 1997). Starting from the fairly recent distinction between object and spatial visual ability, we wanted to explore how this difference will be reflected in the way navigation tasks are performed.

We investigated the relation between object and spatial visual ability and performance in navigation tasks among two groups of students. One group were students majoring in visual art and design at an art academy, as they typically belong to the group of people who are supposed to be object visualizers, according to the studies by Blazhenkova and Kozhevnikov (2010) and Kozhevnikov et al. (2005). The other group were students majoring in a range of engineering studies (industrial automation, electrical engineering, technical computing, construction engineering, and building physics), as they are supposed to be spatial visualizers according to Blazhenkova's and Kozhevnikov's studies.

6.1.4 Research questions

The first research question was if we could replicate the relation between visual-object ability and specialization in visual art/design, and between visual-spatial ability and specialization in engineering/science. The second question was how performance in tasks involving route communication would be affected by visual-spatial abilities. With regard to this second question, our aim was to find differential traces of spatial and object abilities in the way the two groups produce and understand route descriptions and route drawings, while controlling factors we expected to be closely related to the two ability types, like perspective (route versus survey) or route shape. To test this, we designed three tasks. (See Figure 1.)

Controlled route description task

Route information can be described from either a survey perspective – with an extrinsic frame of reference, map-like – or from a route perspective – with an egocentric frame of reference, like ground-level navigation (Taylor & Tversky, 1992; Westerbeek & Maes, 2011). As object visualizers tend to process visual information holistically and to encode and process pictorial properties, we expected them to take a survey perspective more often than spatial visualizers when producing route descriptions based on route maps. We expected the object visualizers to treat the route image as a picture on the plane rather than as a representation of a spatial sequence. Further, as object visualizers are good at identifying global properties characterizing an entire picture (Kozhevnikov et al., 2005), we expected them to refer to the shapes of the routes when describing them more often than spatial visualizers. Also,

we expected them to focus more on discrete entities on the map, like landmarks along the routes, than spatial visualizers, as they are supposed to process object properties rather than spatial relations. On the other hand, we expected spatial visualizers to take a route perspective more often than object visualizers, focusing more on path entities (e.g. streets, patterns), as spatial visualizers tend to encode and process information on spatial relations, and because they will therefore more often than object visualizers process the route image as representing a sequence of steps in space.

Controlled route drawing task

As spatial visualizers are supposed to be good at imagining spatial transformations, we expected them to outperform object visualizers when asked to draw routes on a predefined map, when their drawings are based on descriptions given from a route (as opposed to survey) perspective, as imagining a route description containing subsequent left and right turns requires operations of mental rotation or perspective transformation.

Free map drawing task

When asked to draw a map based on a description including both path entities (e.g. streets) and objects (e.g. landmarks), we expected object visualizers to make more detailed, iconic drawings focusing on objects (landmarks), as they are supposed to generate detailed pictorial images of objects. We expected spatial visualizers to make more schematic drawings focusing on patterns, as they are supposed to generate more schematic images and tend to rely on representations of spatial relations (Kozhevnikov et al., 2002, 2005). As for correctness of drawings, we expected more mistakes to be made in route perspective, and more so by the designers than by the engineers, as route perspective requires mental rotation. Further, we expected the designers to spend more time drawing, as drawing more details would require more time. Further, we expected that drawing papers would be turned around more often when drawing was based on descriptions from route perspective, requiring mental rotation, and more often by the designers than the engineers, as the engineers are supposed to be better at mental rotation operations. Finally, we looked for differences in drawing order of described elements between the two groups and the two conditions.

6.2 Method

Participants

Sixteen bachelor students in art and design (9 male, 7 female, mean age 23) and sixteen bachelor students in engineering (15 male, 1 female, mean age 21) of Avans University of Applied Sciences volunteered to participate in the experiment. Participation was individual.

Design

As to the first research question on the relation between visual ability and profession, participants' visual ability was tested using three tests. The first test was the Object-Spatial Imagery Questionnaire (OSIQ, Dutch translation by Diane Pecher), developed by Blajenkova, Kozhevnikov, & Motes (2006), measuring participants' self-reported object and spatial visual ability. The other two tests were selected from the tests used by Blajenkova et al. to develop the OSIQ. The paper-and-pencil Snowy Pictures Test (Ekstrom, French, & Harman, 1976) is a measure of visual-object ability, measuring people's ability to recognize and identify objects hidden in a degraded picture. (In their study Blajenkova et al. used the Degraded Pictures Test, which is a computerized version of the Snowy Pictures Test.) The Spatial Imagery Test from the Imagery Testing Battery (MMVirtual Design) was used to test participants' visual-spatial abilities. This computerized test consisted of three spatial imagery tasks: Wire Frame (requiring imagining a cube from different perspectives), Figure Rotation and Combination (requiring imagining mentally rotating and combining figures), and Folded Box problems (requiring imagining folding a template into a box).

As to the second research question, performance differences on navigation tasks were measured in the three tasks described above. Profession (designer vs. engineer) was a between participants factor, perspective (survey or route) was a between participants factor in task 2 and 3, divided evenly over the two types of professionals, and route shape (regular or irregular) was a within participants factor in task 1. For each participant, the perspective in task 2 and 3 alternated.

Materials

For the first task, controlled production of route descriptions, we constructed 16 routes. Each route was drawn on the same map grid of square blocks and perpendicular streets, each turning point being flagged by external landmarks in the form of written labels (e.g. 'cafe', 'gas station'). See Figure 1. Half of the routes had the shape of an object (like a square box or stairs), the other half had an irregular shape. Furthermore, the routes systematically varied in terms of length (4 or 6 blocks involved), number of turning points (2 to 5) and route orientation (left-right, top-bottom and vice versa).

For the second task, controlled route drawing, we constructed two route descriptions, one in a survey, one in a route perspective, for each of these 16 route maps. See Figure 1 for an example. These descriptions were recorded in two sets of 16 spoken directions: one set from a survey and one from a route perspective. (See Figure 1 for examples.)

For the third task, map drawing, two descriptions were composed of a small town consisting of a central park surrounded by four streets and five landmarks (e.g. a school, a supermarket), one from a survey and one from a route perspective, of comparable length (137 and 132 words). See Figure 1.

Procedure

Participants took the tests sitting in a quiet room behind a computer, pressing the space bar to proceed through the instructions in task 1, and with pencil and paper for completing tasks 2 and 3. All experiments were audio and video recorded. Each session was taken individually and took about one hour to complete. The order of tasks was the same for all respondents.

Task 1. Each participant was presented with the 16 route images on the computer screen, one by one in random order, and asked to describe each route. Participants were asked to imagine an addressee for their route descriptions who should be able to easily recognize each described route in a set with all route maps.

Task 2. Each participant was presented with spoken descriptions of the 16 routes, one by one in random order, either from a survey or from a route perspective. They were asked to draw each described route (pen on paper) once the description was finished on a predefined map with only a starting point depicted on it.

Task 3. Each participant was presented with a description on the computer screen of a small town, either from a route or from a survey perspective. They were asked to draw a map of the described town, taking as much time and looking back on the screen as often as they needed.

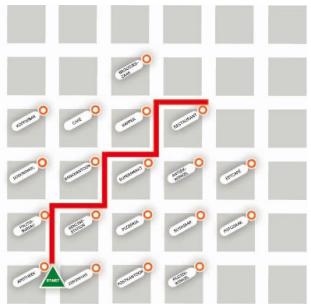
Task 1	Task 2	Task 3
Route description task one of 16 route images	Route drawing task one of 16 route descriptions	Map drawing task part of town description
	<p><i>Survey version:</i> Go one block up. Go one block to the right. Go one block up. Go one block to the right. Go one block up. Go one block to the right.</p> <p><i>Route version:</i> Go right at the next intersection. Go left at the next intersection. Go right at the next intersection. Go left at the next intersection. Go right at the next intersection Go to the next intersection.</p>	<p><i>Survey version:</i> The small village of Etna is built around the city park, which is a square of 400 by 400 meters in the center of the village. The four main streets of Etna run in a square along the park. Along the right side of the park runs Hill Road. [...]</p> <p><i>Route version:</i> You are at the gas station of the small village of Etna, on the corner of Beech Avenue. You walk into Beech Avenue. On your right you see the city park. You walk straight, along the park. After 400 meters you need to turn right in School Street. [...]</p>

Figure 1 Examples of stimuli used in task 1, 2 and 3.

6.3 Results

6.3.1 Visual ability tests

As expected, the designers scored higher on the OSIQ object scale than the engineers ($t(30) = 3.05, p < .005$), while the engineers scored higher on the spatial scale than the designers ($t(20) = 4.23, p < .001$). The designers scored higher than the engineers on the Snowy Pictures Test ($t(30) = 2.03, p < .05$). Surprisingly, the Spatial Imagery Ability Test showed no difference between the two groups, except in executing time: the designers took much more time to complete it than the engineers ($t(22) = 2.05, p < .05$).

Based on these results we conclude that the group of designers can be regarded as representatives of object visualizers and the group of engineers as representatives of spatial visualizers. See Figure 2.

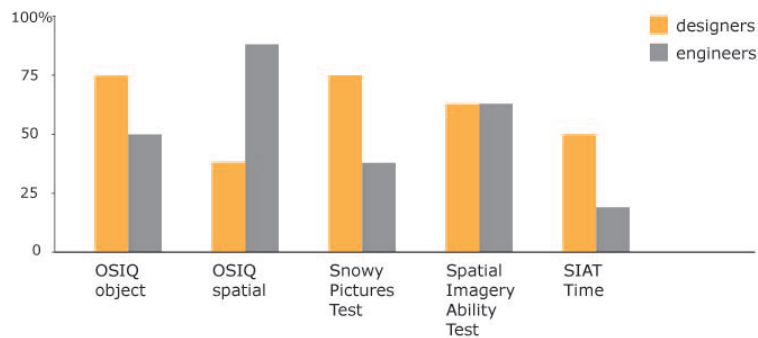


Figure 2. Scores visual ability tests (percentage correct)

6.3.2 Task 1: route description

Analysis

The perspective used was assessed by comparing each description with its depicted route. The expressions 'up' and 'down' were coded as survey perspective. The expressions 'to the left' and 'to the right' were coded as survey when the route went to the left or right from bird's eye perspective and as route when it was to the left or right for someone navigating the route. In cases of doubt, when the route went literally to the left or right but also turned left or right, the expression was neglected. A description was coded as survey or route when two third or more of the coded expressions within the description was labeled as such. Otherwise, the whole description was coded as mixed. Only route action expressions were coded. Descriptive expressions (e.g. descriptions of the starting point on the map) were excluded. Further, we counted all route descriptions in which one or more landmarks were mentioned, as well as all routes in which path entities (street, road) were mentioned. We also counted the total number of references made to either landmarks or path entities by both groups. As for route shapes, we counted the route descriptions containing references to shapes.

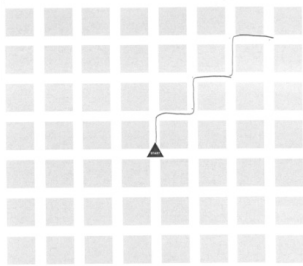

Task 1	Task 2	Task 3
Route description task	Route drawing task	Map drawing task
<p><i>Survey perspective:</i> One block up, one block to the right, one block up, one block to the right, one block up, one block to the right.</p> <p><i>Route perspective:</i> Start at the pharmacy, walk to the police station, there you turn to the right, walk until the gas station, there you turn to the left to the bank, at the bank you turn right to the supermarket, and there you go left to the hairdresser, and at the hairdresser you turn right and then you arrive at the restaurant.</p>		

Figure 3 Examples of results in task 1, 2 and 3.

Results

Almost all participants ($n = 29, 94\%$) showed a preference for using one perspective consistently, meaning they used the same perspective in at least 14 out of 16 route descriptions (see Figure 4). Nine participants consistently used the survey perspective. As expected, more designers than engineers (7 vs. 2) used this perspective. Also according to expectancies, the mean number of routes including discrete entities (landmarks) was higher for the designers than for the engineers ($t(23) = 1.78, p < .05$). Conversely, the mean number of routes including path entities (references to streets) was higher for the engineers than for the designers ($t(17) = 1.96, p < .05$). See Figure 4. Further, the mean number of landmarks mentioned was higher for the designers than for the engineers ($t(25) = 2.33, p < .05$), and the mean number of path entities (streets) mentioned was higher for the engineers than for the designers ($t(29) = 1.75, p < .05$). See Figure 5.

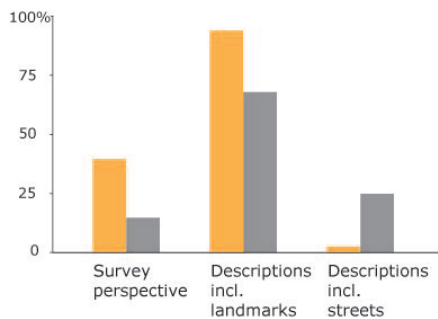


Figure 4. Percentage of route descriptions with survey perspective, and including landmarks and streets

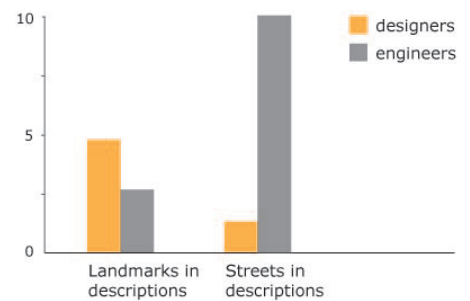


Figure 5. Mean number of landmarks and streets in descriptions

Unexpectedly, visual-object respondents did not refer to the shape of routes more often than visual-spatial respondents. In fact, respondents did not use many references to the shapes of the routes at all: designers referred to shapes in 9% of the route descriptions, engineers did this in 12% of all route descriptions.

6.3.2 Task 2: route drawing

Analysis

Correctness of the route drawings was coded in three categories: incorrect (incorrect turns with or without incorrect number of blocks), pattern correct (correct turns, only one or more blocks too few or too many) and completely correct.

Results

More errors were made in drawings based on directions given from a route (as opposed to survey) perspective ($t(30) = 4.62, p < .001$), as was expected, because route perspective requires operations of mental rotation or perspective transformation. See Figure 6. Scores on the OSIQ spatial scale did correlate with scores on drawing correctness ($r = .42, p < .05$). However, there was no difference between the two groups of professionals. Comparison between participants with scores above average in either spatial visual ability test did not show any difference either.

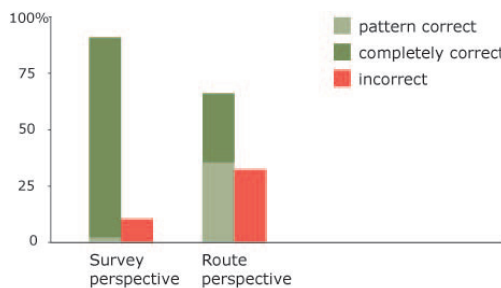


Figure 6. Percentage of correct route drawings

6.3.3 Task 3: map drawing

Analysis

The drawings of landmarks were coded in two levels of detail categories: abstract or iconic. A landmark was coded abstract when it was a simple graphic mark like a blob, a dot or a two-dimensional rectangular. It was coded iconic when it contained graphical elements that resembled one or more characteristics of the landmark (like a roof, a window or door, a supermarket cart, etc.) or was drawn three-dimensionally. The level of detail of street drawings was not coded, as streets offer not many more options than one

or two lines with or without dots in the middle, and it was not always certain whether they consisted of two lines or of one line and the border of the park.

Correctness of maps was restricted to counting the number of correctly placed landmarks.

Reading and drawing times were recorded using a stopwatch while watching the video recordings: time participants took to read the description before they started drawing, and total time reading and drawing.

The number of times participants turned their drawing paper back to the original starting point was counted while watching the videos.

Drawing order was determined by watching the videos and listing all streets and landmarks in the order in which they were drawn and comparing this to the order in which they appeared in the description.

Results

As for level of detail in drawing, the designers drew more iconic landmarks than the engineers, as expected: 27 vs. 14 (designers: $M = 1.63$, $SD = 2.6$; engineers: $M = .88$, $SD = 1.7$).

Almost all landmarks were correctly placed; there was no difference between the groups.

The designers did indeed take more time to draw than the engineers (designers: $M = 168$, $SD = 72$; engineers: $M = 133$, $SD = 48$). When drawing from survey perspective engineers tended to need more reading time than the designers (designers: $M = 161$, $SD = 61$; engineers: $M = 228$, $SD = 137$), while in route perspective they used an equal amount of reading time. See Figure 7.

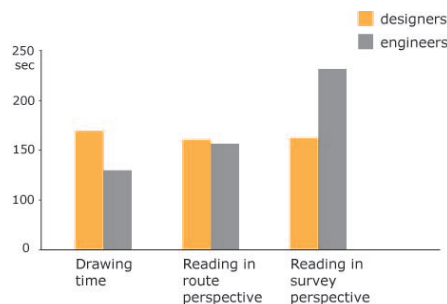


Figure 7. Number of seconds drawing and reading during map drawing

When drawing from route perspective designers tended to turn their drawing paper back to the original starting position more often than the engineers (designers: $M = 1.25$, $SD = 1.04$; engineers: $M = .43$, $SD = .79$).

In drawing order, there was no major difference between the two groups; most participants largely follow the text order. It is notable though, that some designers showed a preference for drawing landmarks first and some engineers for drawing streets first, and none vice versa.

6.4 Discussion

The results from the visual ability tests, specifically the OSIQ, confirm Blahzenkova's and Kozhevnikov's findings that object and spatial imagery are related to specialization in visual arts and engineering respectively. An important and unexpected finding is that designers and engineers scored equal in the Spatial Imagery Ability Test. This result might be accounted for by the fact that the designers took much more time to execute the test, which would mean that lower ability can be compensated by more time and greater effort. Possible gender effects have not been accounted for in this study. It would be valuable to analyze these as well in future studies on this subject.

The production of route descriptions show that differences in visual ability do indeed result in different ways of processing route images. Designers, associated with object imagery, tend to describe them from a survey perspective and focus more on objects (landmarks), while engineers, associated with spatial imagery, tend to describe them more from a route perspective, step-by-step, focusing on paths (street patterns). Unexpectedly, not many participants mentioned the shape of the routes. Recall that we asked respondents to imagine an addressee who should be able to select each described route in a visual array of many different route maps. Apparently, this assumed communicative context was too weak to stimulate respondents to use the shape of the route as a descriptive attribute. It is likely to assume – as a small scale pretest suggested – that a real addressee with a visual array would increase the relevance of shape considerably.

The controlled drawing task resulted in more errors when drawing was based on route (as opposed to survey) descriptions. But the two groups did not differ, probably because the overall complexity level was low, and too subtle to show differences in mental transformation abilities between the groups.

The difference in level of detail in the drawings, and in drawing time, between designers and engineers reflects Blahzenkova and Kozhevnikov's findings that object visualizers tend to process and generate vivid and detailed images. Further, the fact that in task 3 designers tended to turn their drawing paper back to the original position more often than the engineers when drawing from descriptions in route perspective, clearly suggests that the designers have more difficulty with perspective taking transformations (associated with spatial visual ability) than the engineers.

The equal scores between the two groups in the Spatial Imagery Ability Test calls for further analysis of the relation between professions and spatial ability tests. But altogether, the findings in this study show that differences in visual ability lead to differences in processing and producing verbal and visual route information.

Chapter 7

General conclusion and discussion

In this thesis we investigated the main quality criteria for data visualizations for a broad audience, from the perspectives of designers and laypeople. We focused on the way information visualizations for a broad audience are produced, understood, and evaluated by these two target groups. Since this thesis consists of self-contained articles, the findings of the studies are also discussed in the previous chapters. In this chapter we therefore briefly summarize these findings. Then we discuss how these findings answer our research questions, and give suggestions for future research.

7.1 Summary of findings

Chapter 2: Information visualization for a general audience: the designer's perspective

This chapter aimed at identifying designers' criteria for good information visualization for a general audience. How important do they consider clarity and attractiveness? Do they intend to communicate objective information or subjective meaning? And do they have ideas about what makes an information visualization attractive? We collected quotes containing normative expressions with regard to these questions from interviews with designers and design literature. These quotes revealed that designers consider attractiveness and clarity the two main quality criteria for information visualizations for a broad audience, but that clarity is most important. Attractiveness is widely considered important by designers and design literature, but clarity is claimed to be paramount. Further, designers do not feel the need to convey their own truths or opinions, but accept that raw data have to be interpreted for their audiences. Regarding aesthetics, it is striking how little can be found in design literature on what makes visualizations attractive. The quotes from the interviews show that it is difficult for designers to put into words what makes an information visualization attractive.

Chapter 3: Would you prefer pie or cupcakes? Preferences for data visualization designs of professionals and laypeople in graphic design

In chapter 3 we investigated to what extent designers and laypeople share ideas about clarity and attractiveness. The two groups were asked to evaluate information visualizations produced by designers. These designs varied from standard bar and pie graphs to non-standard designs (other constructions), and from lacking to containing pictorial elements. Participants were asked to evaluate the attractiveness, clarity, and overall quality of the designs, to rank the five best and worst designs, and to explain their motives for the rankings. Results show that designers find non-standard constructions the most attractive, whereas laypeople in design find standard constructions the most attractive. The verbal explanations of the rankings reveal that designers prefer designs that are different, whereas laypeople prefer designs they consider clear. Further, designers rate pictorial visualizations higher than abstract ones, whereas the reverse is true for laypeople.

Chapter 4: Graph and chart aesthetics for experts and laypeople in design: The role of familiarity and perceived ease of use

In chapter 4 the relationship between familiarity, usability, and attractiveness of information visualizations was investigated. Does familiarity positively affect perceived usability and attractiveness? And does familiarity affect attractiveness differently for designers and laypeople in design? In study 1, a survey was conducted to assess the perceived attractiveness, familiarity and perceived usability of a series of graphs ranging from familiar to novel. Designers and laypeople in design judged the same graphs most and least familiar, and perceived the same graphs as most and least usable. Familiarity showed to be a strong predictor of perceived ease of use for both target groups. The two groups differed in their attractiveness judgments though: laypeople are attracted to designs they perceive as familiar and usable. Designers are attracted to designs between familiar and novel. In study 2, a different group of participants was asked to use the same graphs in an information retrieval task, and then to rate each graph's attractiveness. Laypeople's judgements of attractiveness remained the same, but the designers were more attracted to familiar designs than in study 1. Correlational analyses suggest that their attractiveness judgments after use were affected not by actual usability, but by perceived usability.

Chapter 5: Reading graphs. The role of length and area in comparing quantities

In chapter 5 we investigated what features affect the usability of popular information visualizations. First, we established which perceptual features are perceived by non-expert users to play a crucial role in everyday usage of graphs. We investigated this by conducting a survey, asking respondents to judge the relative importance of four perceptual features in comparing magnitudes in a series of familiar and more novel graph types: length, position, area, and angle. The results show that non-expert users have different opinions than researchers about the role of the perceptual features under study. First, in all graphs, more than one feature is perceived to play a role in the comparison of magnitudes. Second, in all graphs, either length or area or both play a significantly more important role than the other features. Third, for the majority of graphs area is believed to play the most crucial role. Subsequently, another group of participants was asked to use the same set of graphs in a task reflecting everyday use: comparing one magnitude to another (simple comparison) and comparing one magnitude to a mental summation of two other magnitudes (complex comparison). The results show that performance with length is more accurate and efficient than with area representing quantity, but only in complex comparisons.

Chapter 6: Visual Ability in Navigation Communication

In the final study we investigated the influence of differences in visual ability on the way designers and laypeople produce and process visual information.

Studies have shown that object and spatial visual abilities are distinct types of visual intelligence. Object visualizers – associated with specialization in art and design – rely on ‘visual imagery’ and generate detailed pictorial images of objects. Spatial visualizers – associated with specialization in sciences and engineering – rely on ‘spatial imagery’ and are good at generating schematic images representing spatial relations and imagining spatial transformations. We investigated if and how this difference in visual ability is reflected in performance in specific visual communicative tasks, in this case navigation communication. Participants (designers and engineers) were asked to visualize verbal route information, to describe visual route information, and to draw route maps based on descriptions. Results confirm that designers can be considered object visualizers, and engineers can be considered spatial visualizers. Results further show that designers process route information more often from a bird’s eye perspective, focus more on objects (landmarks) and create more detailed drawings than spatial visualizers. Conversely, engineers process route information more often from a route perspective, focus more on path entities (streets), and create more schematic drawings than object visualizers. Apparently peoples’ innate visual abilities determine in part the way they process and produce (visual) information.

7.2 Conclusions and discussion

The results from each of the five studies as summarized in the previous section contribute to answering several of the research questions we set out in the introduction.

7.2.1 What is the importance of functional and aesthetic criteria in judging visualizations?

We asked ourselves what the main quality criteria are for information visualizations for a general audience. Little is known about how designers think about the relative importance of clarity and aesthetics in information visualization. Similarly, little is known about the way laypeople, as the users of their work, understand and appreciate their designs. In our studies we investigated the role and relative importance of functional and aesthetic criteria from the perspectives of both groups.

Contrary to what some researchers assume (e.g. Kosara, 2007; Vande Moere & Purchase, 2011), the results of the interviews and literature review (chapter 2) show that designers consider clarity the most important criterion, and that they intend to convey information to their audience as objectively and correctly as possible. Attractiveness is considered important, but clarity is paramount.

The results of the study in chapter 3 show that designers and laypeople differ in their opinions about clarity and attractiveness. Designers are attracted

to pictorial rather than abstract designs, whereas the reverse is true for laypeople. The designs that laypeople choose as the best are standard designs, which they consider the clearest. The designs that designers choose as the best are non-standard designs, which they consider not the clearest, but the most attractive, contrary to their statements in the interviews. These results are in line with the results from the study in which designers and laypeople judged the familiarity, perceived usability and attractiveness of information visualizations (chapter 4). These showed that laypeople are attracted to graphs they perceive as familiar and usable, whereas designers are attracted to designs they perceive as more novel and less usable. The results suggest that actual experience with graphs change the designers' preferences: designers who were asked to carry out tasks with the graphs appeared to take into account the perceived usability of graphs more than designers who did not have to carry out tasks with the graphs

Taken together, the results show that laypeople attach more importance to clarity (or perceived usability) than to aesthetics, whereas designers put more emphasis on attractiveness than on clarity. This contrasts with what is claimed in the interviews and design literature, and is in line with assumptions by researchers in information visualization. These results have relevance for the design practice. Assuming that it is a designers' job to create designs that are attractive, but at least also understandable for their intended audience, these results imply that designers would do well to test their designs before publishing. As they acknowledge in the interviews, designers usually test their designs in an informal way, if they test them at all, often among fellow designers. Testing designs among a group of laypeople might yield valuable insights into the way their designs are understood and appreciated by their audiences.

7.2.2 What makes popular information visualizations attractive?

The findings from the interviews and literature review (chapter 2) show that designers assign much importance to aesthetics, but it is hard to find clues in their testimonials as to what factors contribute to aesthetics. Design literature hardly mentions any aesthetic criteria. Answers to interview questions reveal that designers find it very hard to put into words what makes a visualization attractive. This is of course not surprising, since it is difficult in any field to define what makes an artifact aesthetically pleasing.

Based on studies into visual intelligence and on discussions among designers of information visualizations about embellishment, we assumed that the use of pictorial elements could influence attractiveness. The results in chapter 3 showed that designers more so than laypeople appreciate pictorial elements in graphs, which is in line with expectations based on differences between visual-object ability and visual-spatial ability. However, it sounds somewhat paradoxical that a characteristic of graphs that may be seen as being used to make graphs more attractive and popular, is more appreciated by designers than by the broad audience.

Based on evolutionary aesthetics and processing fluency theories, we expected an influence of familiarity on attractiveness as well. This influence was confirmed in the study in which respondents were asked to judge the familiarity, perceived usability and attractiveness of a series of more and less familiar graph designs (chapter 4). Familiarity showed to be a predictor of attractiveness, but stronger for laypeople than for designers. Laypeople are attracted to designs which are familiar and perceived as usable, whereas designers are attracted to more novel (and perceived as less usable) designs.

All things considered, it seems that the perceived usability of visualizations makes them attractive, be it more so for laypeople than for designers. This finding is in line with studies in computer sciences, in which clarity is assumed to be aesthetically pleasing (e.g. Ngo, Teo, & Byrne, 2003). This finding seems to be in line also with theories that hypothesize that processing fluency, resulting from familiarity, is perceived as attractive (e.g. Reber et al., 2004). Based on this theory, one would expect a relationship between performance and attractiveness. But no correlations were found in chapter 4 between performance measures and attractiveness ratings, only between familiarity, attractiveness and *perceived* usability ratings. This suggests that people are attracted to perceived usability, based on familiarity, rather than actual usability. Finally, the results suggest that pictorial elements are attractive, at least for designers. This conclusion is drawn with caution, however. Three of the five designs least appreciated by the designers in this study were also pictorial. This may be caused by the fact that two of these three designs did not show differences in quantities at all, which was not appreciated.

7.2.3

What makes information visualizations usable?

Traditional bar and pie graphs are the most familiar graph types, as our pilot study in chapter 4 shows. Familiarity is a result of frequent exposure, and leads to processing fluency (e.g. Reber et al., 2004). Therefore, one may expect familiar graph types to be the most efficient. Our studies provide mixed results. In chapter 3, performance with standard types – i.e. familiar bar and pie graphs - turned out to be more efficient than performance with non-standard designs, but there are confounding factors that may have influenced these results. Among the non-standard designs in the materials were several designs in which quantities were not accurately readable, which may have affected response times. The performance results in chapter 5 show that performance with the familiar bar and pie graphs is not better than with more novel designs.

Influential graph design research (e.g. Cleveland & McGill, 1984; Spence & Lewandowsky, 1991) focuses on the influence on usability of different perceptual features used to encode quantity. We asked ourselves which features may be responsible for the usability of information visualizations for use by non-expert users in everyday tasks. The results of the survey in chapter 5 show that in each graph more features are perceived to play a role in the comparison of magnitudes, and that in all graphs either length or area, or both,

play a significantly more important role than the other features. The results of a performance task with the graphs consisting of simple and complex comparisons showed that performance with length-types is better in terms of accuracy and efficiency than with area-types, but only in the complex task. Simple comparisons were made as accurately and efficiently with length-types as with area-types.

These results nuance assumptions made by researchers in existing studies regarding which features are responsible for the effectiveness of graphs. Non-expert users' intuitions are not in line with those of researchers in this field, who assume that for example angle is crucial in reading quantity in pie graphs (e.g. Cleveland & McGill, 1984; Simkin & Hastie, 1987), or position along a common scale in bar graphs (e.g. Cleveland & McGill, 1984). In the case of two designs it remained inconclusive if length or area is more important: the divided bar, and the semi-circle. Interestingly, in Cleveland and McGill's (1984) study, claims about judging length being more accurate than judging area are based on experiments in which the divided bar graph is considered to use length to represent quantity, not area. Overall, area was perceived to play a more crucial role than length in the majority of graphs.

The results also show that usability is highly dependent on the type of task. Other types of tasks than the ones we used, may yield different results.

7.2.4 How do designers and laypeople differ in their understanding and aesthetic preferences?

We expected differences between the two target groups for two main reasons. Studies in art have shown differences between experts and novices. And according to studies into visual intelligence, artists and designers have a different kind of visual intelligence than other people, which may cause differences in aesthetic preferences. The results of the studies in this thesis indeed show clear differences between the two groups.

First, the results of the studies in which designers and laypeople are asked to judge the attractiveness and quality of information visualizations (chapters 3, 4) show that laypeople are attracted to familiar designs, which appear clear to them. Designers prefer more novel designs, because they find those attractive. Laypeople attach most importance to clarity, designers to attractiveness. This is in line with studies describing differences between experts (in art) and novices, with novices preferring simple and prototypical stimuli and experts preferring complex and novel stimuli (e.g. McWhinnie, 1968; Reber et al., 2004). These studies focused specifically on appreciation of art works, but our results suggest that similar mechanisms may be at play in design.

Second, the results of chapter 3 show that designers are attracted to pictorial designs and laypeople to abstract ones. In chapter 6 we investigated if differences in visual abilities between designers and laypeople results in different strategies for processing and producing visual information, in this case route information. Results show, among other things, that designers

generate more pictorial and detailed visualizations than engineers. The difference in level of detail in the drawings, and in drawing time, between designers and engineers reflects findings that object visualizers tend to process and generate vivid and detailed images (e.g. Blazhenkova & Kozhevnikov, 2010).

These results show that not only design variables, but also individual differences such as domain or task familiarity (experts vs. novices) and visual abilities influence the way people interact with visual information. More insight in these differences and how they affect the use of visual information may help find visualization solutions that take different preferences and abilities into account.

Above, we discussed the theoretical implications of our findings. But how about their societal relevance? As was mentioned in the introduction, information designers can play an important and responsible role in transforming data into meaningful information for the public. Huge amounts of data need to be visualized to engage a broad public in this information, and to inform them about developments that affect their lives and society. The finding that there are substantial differences between designers' and their audiences' criteria for 'good' information visualizations means that they are facing a communication gap. It is a challenge for the design practice and for design education to think about ways to bridge these differences in aesthetic and functional preferences.

7.3 Future research

Generalizations based on the results of the studies in this thesis should be made with caution. We only tested static 2D information visualizations, and regarding data visualizations, we confined the materials to designs that represent a combination of nominal and quantitative data. These data are quite common in mass media, and we used a wide variety of visualization designs. Yet, many other types of data exist, and visualization techniques, which should be included in future research.

The tasks we used in the information retrieval tasks consisted of simple comparisons (judging if one magnitude is larger or smaller than another) or complex comparisons (judging if a magnitude is larger or smaller than a summation of two other magnitudes). This type of task represents the way information visualizations in mass media are typically used by non-expert users, but different results may result from different kinds of tasks. Further, studies have shown that results may be influenced by an interaction between type of task and construction type. According to several studies (e.g. Simkin & Hastie, 1987) the type of task we used, comparing magnitudes, would be advantageous for graphs in which segments are arranged as separate parts, such as the pie graph. In the design of the materials and the analyses of the results we did not control for such an interaction. It would be interesting for

future research to study interactions between type of task and arrangement with a wider variety of graphs than pie and bar graphs.

The materials we used in the study in chapter 3 showed a wide variety of differences in both construction types and drawing styles. This may have affected preferences in such a way that for example aesthetic preferences for the pictorial mode are influenced by construction type (standard or not) or a drawing style that is appreciated or not. The influence of pictorial elements on attractiveness could better be tested by using materials in which this variable (pictorial elements) is varied in a more controlled manner, with less variation in pictorial style and graph constructions. Similarly, usability results regarding the effect of being standard or non-standard in chapter 3 may have been confounded by the fact that some non-standard designs did not allow accurate readings of the data. The effect of this variable – standard (familiar) vs. non-standard (novel) - could better be studied by using a more controlled set of materials as well.

Our studies show that it is worthwhile to investigate which features are actually perceived as representing quantity in graphs. We used subjective, evaluative measures for this, but it would be interesting to do this by using more objective methods such as eye tracking. Further, performance with those features is probably largely dependent on the type of task. Clearly, our results only apply in the specific type of task we used, involving comparison of relative magnitudes. It would be interesting to investigate the effect of perceptual features in other cognitive tasks as well, again using a wider variety of graph designs than the commonly used bar, divided bar, and pie graphs.

When investigating the relationship between aesthetics and usability, we only measured correlations between the single items 'attractiveness' and '(perceived) usability'. But in fact, both aesthetics and usability are complex constructs. It would be worthwhile to further investigate the relationship between aesthetics and usability by using multi-level scales, which do justice to the multifaceted nature of these notions. In the field of HCI interesting studies have been done attempting to reveal such factors (Lavie & Tractinsky, 2004; Tuch, Roth, Hornbaek, Opwis, & Bargas-Avila, 2012; Hassenzahl, 2004). An interesting direction for future research would be to develop measurement scales that are appropriate for information visualization for a general audience.

Regarding relationships between aesthetics and usability, it would also be interesting to further investigate the relationship between perceived usability, actual usability, and attractiveness. Existing studies have shown positive relationships between perceived beauty and perceived usability, which have led researchers to conclude that perceived beauty leads to perceived usability (e.g. Tractinsky, 2000). However, our studies showed that participants differ in opinions about attractiveness, but have similar perceptions of usability, most probably based on familiarity, of visualizations. This could mean that the relationship between attractiveness and perceived usability is reversed,

meaning that people are attracted to things they perceive as usable, instead of assuming that things they find attractive are usable. This relationship could be further investigated in experiments in which actual usability and perceptions of usability are measured more systematically in relationship with assessments of attractiveness before and after use.

Differences in visual ability seem to have implications for the way people produce and process information visualizations. Our study into such implications was only focused on the production and processing of route information. As Blazhenkova and Kozhevnikov (2010) point out, the role of visual-object ability is increasing in a wide range of professions and in everyday life, as the use of rapidly presented visual stimuli, which call for quick, holistic processing, are prevailing in contemporary media. It would therefore be useful and relevant to further investigate the implications of differences in visual ability also in other forms of visual communication.

There are many ways in which scientific research could contribute further to insights into design practice, by doing the kind of experimental research we conducted in this thesis, by conducting in-depth interviews with designers about specific design cases, by observing designers at work while interviewing them about their reasoning, or by examining and comparing a body of design cases and the contexts in and the users for which they were created. Designers may benefit from the insights that studies into the graphic design practice provide, because they enable them to move from solving one unique case after another to the use of explanatory principles and solutions for similar kinds of problems (Friedman, 2003). Both scholars and practitioners involved in information visualization for broad audiences, could benefit from insights into how information can be visualized in ways that are both understandable and appealing.

References

- Aginski, V., Harris, C., Rensink, R., & Beusmans, J. (1997). Two Strategies for Learning a Route in a Driving Simulator. *Journal of Environmental Psychology* 17, 317-331.
- Aharon, I., Etcoff, N., Ariely, D., et al. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron*, 32(3), 537-551.
- Allen, G.L., Kirasic, K.C., & Dobson, S.H. (1996). Predicting Environmental Learning from Spatial Abilities: an indirect route. *Intelligence* 22, 327-355.
- Arnheim, R. (1969). *Visual Thinking*. Berkeley: University of California Press.
- Aspinwall, L., Shaw, K. L., & Presmeg, N. (1997). Uncontrollable mental imagery: Graphical connections between a function and its derivative. *Educational Studies in Mathematics*, 33, 301-317.
- Baird, J.C. (1970). *Psychophysical analysis of visual space*. New York: Pergamon Press.
- Bateman, S., Mandryk, R.L., Gutwin, C., Genest, A., McDine, D., & Brooks, C. (2010). Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts. In *ACM Conference on Human Factors in Computing Systems (CHI 2010)*, pp. 2573-2582. Atlanta, GA, USA.
- Berlyne, D.E. (1971). *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts.
- Bertin, J. (1983 / 2011). *Semiology of graphics: diagrams networks maps*. Redlands, CA: Esri Press.
- Blackwell, A. & Engelhardt, Y. (2002). A Meta-Taxonomy for Diagram Research. In: Anderson, M., Meyer, B., Olivier, P. (eds.) (2002). *Diagrammatic Representation and Reasoning*, pp. 47-64. London: Springer.
- Blajenkova, O., Kozhevnikov, M., & Motes, M. A. (2006). Object-spatial imagery: A new self-report imagery questionnaire. *Applied Cognitive Psychology*, 20(2), 239-263.
- Blasko, D.G. & Connine, C.M. (1993). Effects of Familiarity and Aptness on Metaphor Processing. *Journal of Experimental Psychology: learning, Memory, and Cognition*, 19(2), 295-308.
- Blazhenkova, O, & Kozhevnikov, M. (2010). Visual-object ability: A new dimension of non-verbal intelligence. *Cognition*, 117, 276-301.
- Bornstein, R.F., Kale, A.R., & Cornell, K.R. (1990). Boredom as a limiting condition on the mere exposure effect. *Journal of Personality and Social Psychology*, 58, 791-800.
- Bourdieu, P. (1987). *Distinction: A social critique of the judgment of taste*. Cambridge, MA: Harvard University Press (original work published 1979).
- Brückner, H. (2004). *Informationen Gestalten*. Bremen: Hauschild.
- Buhrmester, M., Kwang, T., & Gosling, S.D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science*, (6)3, 3-5.
- Burkhard, R.A., Andrienko, G., Andrienko, N., et al. (2007). Visualization Summit 2007: ten research goals for 2010. *Information Visualization*, 6, 169-188.
- Cairo, A. (2013). *The functional art: an introduction to information graphics and visualization*. Berkeley, CA: New Riders.

- Card, S.K., MacKinlay, J.D., & Shneiderman, B (eds.) (1999). *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers.
- Card, S.K., & MacKinlay, J.(1997). The Structure of the Information Visualization Design Space. *Proceedings. IEEE Symposium on Information Visualization, 1997*.
- Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied, 4*, 75-100.
- Carswell, M. (1992). Choosing specifiers: An evaluation of the basic tasks model of graphical perception. *Human Factors, 34*, 535-554.
- Cawthon, N. & Vande Moere, A. (2007). The Effect of Aesthetic on the Usability of Data Visualization. In *Proceedings of the 11th International Conference Information Visualization (IV'07)*, Zürich, Switzerland, 4-6 July 2007, pp.631-636. New York, NY, USA: IEEE.
- Chater, N. (1999). The search for simplicity: a fundamental cognitive principle? *The Quarterly Journal of Experimental Psychology, 52A(2)*, 273-302.
- Chen, C. (2005). Top 10 Unsolved Information Visualization Problems. *IEEE Computer Graphics and Applications, (25)4*, 12-16.
- Cleveland, W.S. & McGill, R. (1984). Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistics Association, (79)387*, 531-554.
- Cleveland, W.S. (1984). Graphs in Scientific Publications. *The American Statistician, 38(4)*.
- Cleveland, W.S., (1985). *The Elements of Graphing Data*. Wadsworth, Monterey, CA.
- Cross, N. (1982). Designerly Ways of Knowing. *Design Studies, 3(4)*, 221-227.
- Descartes, R. (1650). *Musicae Compendium*. Utrecht: van Zi; II & van Ackerodijk.
- Ekstrom, R., French, J., & Harman, H. (1976). *Manual for kit of factor referenced cognitive tests*. Princeton, NJ: Educational Testing Services.
- Engelhardt, Y. (2002). *The Language of Graphics - A framework for the analysis of syntax and meaning in maps, charts and diagrams*. Amsterdam: ILLC, UvA.
- Fabrikant, S. I., Rebich-Hespanha, S., & Hegarty, M. (2010). Cognitively inspired and perceptually salient graphic displays for efficient inference making. *Annals of the Association of American Geographers, 100*, 13-29.
- Fabrikant, S.I, Christophe, S., Papastefanou, G. et al. (2012). Emotional response to map design aesthetics. In *Proceedings of GIScience Conference 2012*, Columbus, Ohio, USA, 18-21 September 2012. Berlin: Springer.
- Farah, M. J., Hammond, K. M., Levine, D. N., & Calvanio, R. (1988). Visual and spatial mental imagery: Dissociable systems of representations. *Cognitive Psychology, 20*, 439-462.
- Few, S. (2004). *Show me the numbers: designing tables and graphs to enlighten*. Oakland, CA: Analytics Press.

- Frascara, J. (2004). *Communication Design - principles, methods, and practice*. New York: Allworth Press.
- Freedman, E.G., & Shah, P. (2002). Toward a model of knowledge-based graph comprehension. In: *Diagrammatic representation and inference*. Springer.
- Friedman, K. (2003). Theory construction in design research: criteria, approaches, and methods. *Design Studies*, 24(6), 507-522.
- Galea, L.A.M., & Kimura, D. (1993). Sex differences in route learning. *Personality and Individual Differences*, 14 (1), 53-65.
- Gaver, W.W., Beaver, J., & Benford, S. (2003). Ambiguity as a resource for design. In: *Human Factors in Computing Systems (CHI)*. Ft. Lauderdale, Florida, USA, ACM Press, pp. 233-240.
- Gog, T. van, & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, 20, 95-99.
- Goldberg, J., & Helfman, J. (2011). Eye tracking for visualization evaluation: Reading values on linear versus radial graphs. *Information visualization*, 10(3), 182-195.
- Gombrich, E.H. (1995). *The story of art*. London: Phaidon.
- Gammel, L., Tory, M., & Storey, M-A. (2010). How Information Visualization Novices Construct Visualizations. *IEEE Transactions on Visual Computer Graphics*, (15)6.
- Harris, R. (1999). *Information graphics: a comprehensive illustrated reference*. Atlanta, GA: Management Graphics.
- Hassenzahl, M. (2004). The Interplay of Beauty, Goodness, and Usability in Interactive Products. *Human-Computer Interactions*, 19, 319-349.
- Healey, C.G., Tateosian, L., Enns, J.T., & Remple, M. (2004). Perceptually based brush strokes for nonphotorealistic visualization. *ACM Transactions on Graphics*, 23(1), 64-96.
- Heer, J., Kong, N., & Agrawala, M. (2009). Sizing the horizon: the effects of chart size and layering on the graphical perception of time series visualizations. In *Proceedings of the 27th international conference on Human factors in computing systems*, pp. 1303-1312. New York, NY, USA: ACM.
- Hegarty, M. (2011). The Cognitive Science of Visual-Spatial Displays: Implications for Design. *Topics in Cognitive Science*, 3(3), 446-474.
- Hegarty, M., Montello, D.R., Richardson, A.E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, 34, 151-176.
- Hegarty, M. & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology*, 91(4), 684-689.
- Hegarty, M. & Waller, D. (2003). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32, 175-191.
- Hegarty, M. & Waller, D. (2006). Individual differences in spatial abilities. In: A. Miyake & P. Shah (Eds.), *The handbook of visuospatial thinking*. Cambridge: Cambridge University Press.

- Hekkert, P. & Leder, H. (2008). Product Aesthetics. In: Schifferstein, H.N. & Hekkert, P. (eds.) *Product Experience*. Amsterdam: Elsevier, pp. 259-285.
- Hekkert, P. (2006). Design aesthetics: principles of pleasure in design. *Psychological Science*, 48, 157-172.
- Hekkert, P., Snelders, D., & Van Wieringen, P.C.W. (2003). 'Most advanced, yet acceptable': Typicality and novelty as joint predictors of aesthetic preference in industrial design. *British Journal of Psychology*, 94, 111-124.
- Herdeg, W. (ed.) (1981). *Diagrams: the graphic visualization of abstract data*. Zürich: Graphis Press.
- Hollands, J. G., & Spence, I. (1998). Judging proportion with graphs: The summation model. *Applied Cognitive Psychology*, 12, 173-190.
- Holmes, N. (2006). *Nigel Holmes on Information Design*. New York: Jorge Pinto Books.
- Holmes, N. (2005). *Wordless Diagrams*. New York: Bloomsbury.
- Hutcheson, F. (1725). *An inquiry into the original of our ideas of beauty and virtue*. London: Darby.
- IIID Japan (ed.) (2005). *Information Design Source Book*. Basel: Birkhauser-Publishers for Architecture.
- Imagery Testing Battery (Version 1.0) [Computer Software]. Newark, NJ:MMVirtual Design, LLC: www.mmvirtualdesign.com.
- Inbar, O., Tractinsky, N., & Meyer J. (2007). Minimalism in Information Visualization - Attitudes towards Maximizing the Data-Ink Ratio. In: *Proceedings of the 14th European conference on Cognitive ergonomics (ECCE'07)*. pp. 185-188. New York: ACM.
- Jacobson, R.E. (ed.) (1999). *Information design*. Cambridge, MA: MIT Press.
- Judelman, G. (2004). Aesthetics and Inspiration for Visualization Design: Bridging the Gap between Art and Science. In: *Proceedings of the eighth International Conference on Information Visualisation (IV'04)*, London, UK, 14-16 July 2004, pp.245-250. New York, NY, USA: IEEE.
- Kamvar, S. and Harris, J. (2009). *We Feel Fine: An Almanac of Human Emotion*. Scribner.
- Klanten, R. (ed.) (2008). *Data Flow - Visualising Information in Graphic Design*. Berlin: Gestalten.
- Klanten, R., Ehmann, S., Bourquin, N., & Tissot, T. (Eds.) (2010). *Data Flow 2 - Visualising Information in Graphic Design*. Berlin: Gestalten.
- Klanten, R. (ed.) (2011) *Visual Storytelling: inspiring a new visual language*. Berlin: Gestalten.
- Kosara, R. (2007). Visualization Criticism - The Missing Link Between Information Visualization and Art. In *Proceedings of the 11th International Conference on Information Visualisation (IV'07)*, Zürich, Switzerland, 4-6 July 2007, pp.631-636. New York, NY, USA: IEEE.
- Kosslyn, S. M. (1994). *Elements of graph design*. New York: W. H. Freeman.
- Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge, MA, US: The MIT Press.

- Kosslyn, S. M., & Koenig, O. (1992). *Wet mind: The new cognitive neuroscience*. New York, NY, US: Free Press.
- Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, 29 (5), 745-756.
- Kozhevnikov, M., Hegarty, M., & Mayer, R. (1999). *Students' use of Imagery in solving qualitative problems in kinematics*. UC Santa Barbara.
- Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002). Revising the visualizer-verbalizer dimension: Evidence for two types of visualizers. *Cognition & Instruction*, 20, 47-78.
- Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory & Cognition*, 33, 710-726.
- Kozhevnikov, M., Motes, M., & Hegarty, M. (2007). Spatial visualization in physics problem solving. *Cognitive Science*, 31, 549-579.
- Kozhevnikov, M., Motes, M.A., Rasch, B., & Blajenkova, O. (2006). Perspective-Taking vs. Mental Rotation Transformations and How They Predict Spatial Navigation Performance. *Applied Cognitive Psychology*, 20, 397-417.
- Kurosu, M. & Kashimura, K. (1995). Apparent usability vs. inherent usability. *CHI '95 Conference Companion*, Denver, Colorado, USA, 7-11 May 1995; 292-293. New York: ACM.
- Lakoff, G. and Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Lau, A. & Vande Moere, A. (2007). Towards a Model of Information Aesthetic Visualization. In *Proceedings of the 11th International Conference on Information Visualization (IV'07)*, Zürich, Switzerland, 4-6 July 2007, pp.87-92. New York, NY, USA: IEEE.
- Lavie, T. & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of Web sites. *International Journal of Human-Computer Studies* 60, 269-298.
- Levy, E., Zacks, J., Tversky, B., & Schiano, D. (1996). Gratuitous graphics? Putting Preferences in Perspective. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '96)*, pp. 42-49. New York: ACM.
- Lima, M. (2011). *Visual complexity: Mapping patterns of information*. New York: Princeton Architectural Press.
- Lindauer, M. S. (1983). Imagery and the arts. In: A. A. Sheikh (Ed.), *Imagery: Current theory, research, and application*, pp. 468-506. New York: Wiley.
- Lloyd, K.B. & Jankowski, D.J. (1999). A cognitive information processing and information theory approach to diagram clarity: A synthesis and experimental investigation. *The Journal of Systems and Software*, 45, 203-214.
- Lohman, D. (1979). Spatial ability: A review and reanalysis of the correlational literature. *Tech.Rep. No. 8*. California, Stanford University, Aptitude Research Project, School of Education.

- Lohse, G., Biolsi, K., Walker, N., & Rueler, H. (1994). A classification of visual representations. *Communications of the ACM*, 37(12), 36-49.
- MacDonald-Ross, M. (1977). How Numbers Are Shown: A view of research on the Presentation of Quantitative Data in Texts. *AV Communication Review*, 25, 359-409.
- Mackinlay, J. D. (1986). Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*, 5, 110-141.
- Maeda, J. (1999). *Design by numbers*. Cambridge, MA: MIT Press.
- Mayer, R.E., & Massa, L.J. (2003). Three Facets of Visual and Verbal Learners: Cognitive Ability, Cognitive Style, and Learning Preference. *Journal of Educational Psychology*, 95 (4), 833-846.
- McCandless, D. (2009). *Information is beautiful*. London: Collins.
- McGee, M. (1979). Human Spatial Abilities: Psychometric Studies and Environmental, Genetic, Hormonal, and Neurological Influences. *Psychological Bulletin*, 86 (5), 889-918.
- McWhinnie, H.J. (1968). A review of research on aesthetic measure. *Acta Psychologica*, 28, 363-375.
- Miller, A. I. (1996). *Insights of genius imagery and creativity in science and art*. Cambridge, Mass: MIT Press.
- Mijksenaar, P. (1997). *Visual Function: an introduction to Information Design*. Rotterdam: 010 Publishers.
- Moffat, S.D., Hampson, E., & Hatzipantelis, M. (1998) Navigation in a 'Virtual' Maze: Sex Differences and Correlation With Psychometric Measures of Spatial Ability in Humans. *Evolution and Human Behavior* 19, 73-87.
- Neurath, O. (2010). *From hieroglyphics to Isotype: a visual autobiography*. London: Hyphen.
- Ngo, D., Teo, L., & Byrne, J. (2003). Modelling interface aesthetics. *Information Science*, 152(25), 25-46.
- Norman, D. (2004). *Emotional Design. Why we love (or hate) everyday things*. New York: Basic Books.
- Norman, D. (2007). Simplicity is highly overrated. *Interactions*, 14(2). New York, NY: ACM.
- Paivio, A. (1971). *Imagery and verbal processes*. Oxford: Holt, Rinehart & Winston.
- Palmer, S.E., Schloss, K.B., & Sammartino, J. (2013). Visual Aesthetics and Human Preference. *Annual Review of Psychology*, 64, 77-107.
- Pellegrino, J.W., Mumaw, R.J., & Shute, V.J. (1985). Analyses of spatial aptitude and expertise. In: S.E. Embretson (Ed.), *Test design: developments in psychology and psychometrics*. Orlando, FL: Academic Press.
- Polanyi, M. (1966). *The tacit dimension*. Garden City, NY: Doubleday.
- Pousman, Z., Stasko, J., & Mateas, M. (2007). Casual Information Visualization: Depictions of Data in Everyday Life. *IEEE Transactions on Visualization and Computer Graphics* 13(6), 1145-1152.

- Quispel, A. & Maes, A. (2014). Would you prefer pie or cupcakes? Preferences for data visualization designs of professionals and laypeople in design. *Journal of Visual Languages & Computing*, 25(2), 107-116.
- Quispel, A., Maes, A., and Schilperoord, J. (in press, first published online September 2015). Graph and chart aesthetics for experts and laymen in design: The role of familiarity and perceived ease of use. *Information Visualization*.
- Reber, R., Schwartz, N., & Winkielman, P. (2004). Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience? *Personality and Social Psychological Review*, 8(4), 364-382.
- Rheingans, P. & Ebert, D. (2001). Volume illustration: Nonphotorealistic rendering of volume models. *IEEE Transactions on Visual Computer Graphics*; 7(3), 253-264.
- Richards, C.J. (2002). The Fundamental design variables of diagramming. In: Anderson, M., Meyer, B., & Olivier, P. (Eds) (2002). *Diagrammatic representation and reasoning* (pp. 85-102). London: Springer.
- Richardson, A. (1977). Verbalizer-visualizer: a cognitive style dimension. *Journal of Mental Imagery*, 1, 109-125.
- Rogers, S. (2012), How data changed journalism. In: Rendgen, S. & Wiedemann, J. (eds.) (2012). *Information Graphics*. Köln: Taschen.
- Rosenberg, H. S. (1987). Visual artists and imagery. *Imagination, Cognition, and Personality*, 7, 77-93.
- Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, 45, 115-143.
- Schön, D.A. (1983). *The Reflective Practitioner. How Professionals Think in Action*. London: Templer-Smith.
- Schubert, T.W. (2005). Your Highness: Vertical Positions as Perceptual Symbols of Power. *Journal of Personality and Social Psychology*, 89(1), 1-20.
- Shah, P. & Carpenter, P. (1995). Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General*, 124, 337-370.
- Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualization, presented at *Visual Languages 96*, pp. 336-343.
- Simkin, D. K., & Hastie, R. (1987). An information-processing analysis of graph perception. *Journal of the American Statistical Association*, 82, 454-465.
- Skog, T., Ljungblad, S., & Holmquist, L.E. (2003). Between Aesthetics and Utility: Designing Ambient Information Visualizations. In *Proceedings of InfoVis 2003*, pp 233-240.
- Spence, I. & Lewandowsky, S. (1991). Displaying proportions and percentages. *Applied Cognitive Psychology*, 5, 61-77.
- Spence, R. (2001). *Information Visualisation*. New York: ACM Press.
- Storkerson, P.K. (2006). Communication Research: Theory, Empirical Studies, and Results. In: Bennet, A. (ed.). *Design Studies - theory and research in graphic design*. New York: Princeton Architectural Press.

- Taylor, H.A. & Tversky, B. (1992). Spatial Mental Models Derived from Survey and Route Descriptions. *Journal of Memory and Language* 31, 261-292.
- Tractinsky, N. & Meyer, J. (1999). Chartjunk or goldgraph? Effects of presentation objectives and content desirability on information presentation. *MIS Quarterly*, 23(3), 397-420.
- Tractinsky, N. (2004). Toward the Study of Aesthetics in Information Technology. *ICIS 2004 Proceedings*. Paper 62.
- Tractinsky, N., Katz, A.S., & Ikar, D. (2000). What is beautiful is usable. *Interacting with computers*, 13, 127-145.
- Tuch, A.N., Roth, S.P., Hornbaek, K., Opwis, K., & Bargas-Avila, J. (2012). Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI. *Computers in Human Behavior*, 28, 1596-1607.
- Tufte, E. R. (1983). *The Visual Display of Quantitative Information*. Cheshire, Connecticut: Graphics Press.
- Tufte, E. R. (1990). *Envisioning Information*. Cheshire, Connecticut: Graphics Press.
- Tufte, E. R. (1997). *Visual Explanations*. Cheshire, Connecticut: Graphics Press.
- Tufte, E. R. (2001). *The Visual Display of Quantitative Information*. Cheshire, Connecticut: Graphics Press.
- Tufte, E.R. (2006). *Beautiful Evidence*. Cheshire, Connecticut: Graphics Press.
- Tversky, B. & Kessel, A. (2006). Using Gestures and Diagrams. *Proceedings of Sig2 Bi-annual meeting Text and Graphics Comprehension, 2006*.
- Tversky, B. (2011). Visualizing Thought. *Topics in Cognitive Science* 3, 4990-535.
- Tversky, B., Kugelmass, S., and Winter, A. (1991). Cross-Cultural and Developmental Trends in Graphic Productions. *Cognitive Psychology*, 23, 515-557.
- Vande Moere, A. & Purchase, H. (2011). On the role of design in information visualization. *Information Visualization*, 10(4), 356-371.
- Vande Moere, A. (2005). Form Follows Data: the symbiosis between Design & Information Visualization. In: *International Conference on Computer-Aided Architectural Design (CAADFuture)*, pp. 167-176, Vienna, Austria, Dordrecht: Springer.
- Ware, C. (2004). *Information Visualization. Perception for design*. San Francisco, CA: Morgan Kaufman.
- Ware, C. (2008). *Visual thinking for design*. Burlington, MA: Morgan Kaufman.
- Ware, C., Purchase, H., Colpoys, L., et al. (2002). Cognitive measurements of graph aesthetics. *Information Visualization*, 1, 03-10.
- Westerbeek, H. G. W. & Maes, A. (2011). Referential scope and visual clutter in navigation tasks. In: K. Van Deemter, A. Gatt, R. Van Gompel, & E. Krahmer (Eds.), *Proceedings of PRE-CogSci 2011*, Boston.
- Wildbur, P. (1989). *Information Graphics*. Houten: Gaade.
- Wildbur, P. & Burke, M. (1998). *Information Graphics - Innovative Solutions in Contemporary Design*. London: Thames and Hudson.
- Woolman, M. (2002). *Digital Information Graphics*. London: Thames & Hudson.

- Wurman, R.S. [website] (accessed 20 January 2016). *Information Anxiety: Towards understanding* (2012). <http://scenariojournal.com/article/richard-wurman/>
- Yau, N. (2011). *Visualize This. The Flowing Data Guide to Design, Visualization, and Statistics*. Indianapolis: Wiley.
- Zacks, J. & Tversky, B. (1999). Bars and Lines: A Study of Graphic Communication. *Memory & Cognition*, 27(6), 1073-1079.
- Zacks, J., Levy, E., Tversky, B., & Schiano, D. (2002). Graphs in print. In: Anderson, M., Meyer, B., & Olivier, P. (Eds.). *Diagrammatic representation and reasoning*. (pp. 187-206). London: Springer Verlag.
- Zacks, J.M., Mires, J., Tversky, B., & Hazeltine, E. (2002). Mental spatial transformations of objects and perspective. *Spatial Cognition and Computation* 2, 315-332.
- Zajonc, R.B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monograph Supplements*, 9(2 Pt 2), 1-27.
- Zajonc, R.B. (1984). On the primacy of affect. *American Journal of Psychology*, 39, 117-123.

Summary

Thus far, the study of Information visualization mostly focused on visualizations that are meant to facilitate an accurate and efficient reading of the visualized data. Numerous studies have investigated features that enhance their effectiveness for this purpose. Far less is known about what makes ‘good’ information visualizations for a broad audience. ‘Popular’ information visualizations are increasingly published in mass media, showing a variety of familiar, but also novel visualization techniques, and not only plain abstract, but also more embellished, illustrative types. What criteria do designers use for such visualizations? To what extent do they consider adequacy, understandability, and attractiveness important? And what is the effect of using novel visualization techniques and pictorial elements on their understandability and attractiveness? Similarly, little is known about the way the general public understands and appreciates these visualizations. To what extent do they share opinions with the designers about the importance of clarity and attractiveness, and about what makes a visualization attractive? In this thesis, we investigated information visualizations for a broad audience: how are they produced, understood, and evaluated by their producers, design experts, and by their audience, laypeople in design? What are the main criteria, and (how) do these criteria differ for designers and laypeople?

We addressed four main questions, which we answered by conducting several studies, as reported on in chapters 2 – 6.

In chapter 2 we identified designers’ criteria for good information visualization for a general audience, by conducting interviews with professional designers, and by reviewing design literature that is recommended and frequently consulted by designers. In chapter 3 we investigated to what extent graphic designers and their audience of laypeople in design share ideas about the clarity and attractiveness of information visualizations. In chapter 4, the influence of familiarity on (perceived) ease of use and on attractiveness of information visualizations was investigated. In chapter 5 we investigated which perceptual features are perceived as representing quantity, and how these features – such as length or area – affect the usability of a series of more and less novel information visualizations. And in chapter 6 we studied how differences in people’s visual abilities are reflected in the way they perform in tasks requiring the use of visual communication.

Below we summarize how the findings of these studies answer the main research questions.

What is the importance of functional and aesthetic criteria in judging visualizations?

The interviews and literature review in chapter 2 suggest that designers consider clarity the most important criterion, and that they intend to convey

information to their audience as objectively and correctly as possible. Attractiveness is considered important, but clarity is paramount. However, the results from the experiments in chapters 3 and 4 show that designers put more emphasis on attractiveness than laypeople, whereas laypeople assign more importance to clarity (or perceived usability) than designers.

What makes popular information visualizations attractive?

Design literature hardly mentions any aesthetic criteria, and interviewees find it very hard to put into words what makes a visualization attractive (chapter 2). This is not surprising, since it is difficult in any field to define what makes an artifact aesthetically pleasing.

We expected an influence of familiarity on attractiveness. Familiarity indeed showed to be a predictor of attractiveness, but stronger for laypeople than for designers (chapter 4). Laypeople are attracted to designs which are familiar and perceived as usable, whereas designers are attracted to more novel (and perceived as less usable) designs. No relationships were found between performance measures and attractiveness ratings, only between familiarity, attractiveness and *perceived* usability ratings. This suggests that people are attracted to perceived usability, based on familiarity, rather than actual usability.

The results of the studies further suggest that pictorial elements are attractive, but more so for designers than for laypeople (chapter 3). This conclusion is drawn with caution, however, since the composition of the stimuli may have biased these results.

What makes information visualizations usable?

We expected familiar information visualizations (bar and pie graphs) to be the most efficient. Our studies provide mixed results: in some they are more efficient than novel types (chapter 3), in others performance with familiar types is not better than with novel types (chapter 4).

We asked ourselves which features may be responsible for the usability of information visualizations for use by non-expert users in everyday tasks (chapter 5). The results of a survey show that non-expert intuitions are not in line with experts' assumptions with this regard. In all graphs either length or area, or both, are perceived to play a significantly more important role than the other features (and never angle, or position along a common scale). The results of a performance task with the visualizations show that performance with length-types is better in terms of accuracy and efficiency than with area-types, but only in complex comparison tasks. Simple comparisons were made as accurately and efficiently with length-types as with area-types. These performance results show that usability is highly dependent on the type of task.

How do designers and laypeople differ in their understanding and aesthetic preferences?

The results of our studies show that laypeople are attracted to familiar designs, which appear clear to them, whereas designers prefer more novel designs, because they find those attractive (chapters 3, 4). Laypeople assign most importance to clarity, designers to attractiveness. Second, results show that designers are attracted to pictorial designs and laypeople to abstract ones (chapter 3). And third, designers and laypeople appear to use different strategies for processing and producing visual information (chapter 6). These results show that not only design variables, but also individual differences (e.g. in experience) and visual abilities influence the way people interact with visual information.

Investigating information visualization is relevant for a number of reasons. Enormous amounts of data need to be visualized for the general public. Information design and designers are increasingly important, but little is known about design practice. This thesis contributes to a better understanding of the designers' practices, of the quality criteria used by designers and their audiences, and of design characteristics determining the usability and attractiveness of such information visualizations.

Acknowledgements

I would like to thank:

Fons Maes, my supervisor, for his guidance, meticulous reviews of manuscripts, good-humored patience when I became impatient sometimes, inspiring discussions about research and design, and for supporting me by granting me access to not only all university facilities, but also to colloquia, PhD meetings and social events.

Joost Schilperoord, who became my supervisor later in the process, for his guidance, interest, help with statistics, manuscript reviews, and encouragement.

Avans University of Applied Sciences and the Centre of Expertise Art & Design of Avans for financing this PhD research, and Rens Holslag, director of the Centre of Expertise, for also helping me with good advice in many stages of the project.

The management of AKVSt.Joost, academy of art and design of Avans University, for giving me the opportunity to work on this research for four years, letting me put other tasks aside for a while.

My colleagues at AKVSt.Joost for their interest and support, especially my teammates Wim le Mair, Michiel van Opstal, Marieke Mein and Jeroen van de Korput, who remind me regularly that there is more in life than work, such as having drinks and dinners with them after work.

The TiCC colleagues at Tilburg University for their interest and their hospitality in research meetings as well as social events, especially Anja Arts for sharing her room with me (and Portugese adventures and other good times), Jacqueline Dake for helping me with several technical issues, and 'minisig' colleagues Lisanne van Weelden, Marije van Amelsvoort, and Hans Westerbeek for reviewing articles and discussing ideas for experiments.

Information designers Erik van Gameren, Daniel Gross, Remy Jon-Ming, Joris Maltha, Dimitri Nieuwenhuizen, Frédéric Ruys, Yassine Salihine, Karin Schwandt, Eugene Tjoa, and Jan Willem Tulp for their cooperation in interviews about their design practice.

Carel Fransen, who created many of the graphs that I used as stimuli in my experiments.

Irina Shapiro for sharing thoughts about design studies which I could not realize within the time frame of this PhD research, but which we can perhaps conduct in the future.

The many students of the bachelor and master courses in graphic design of AKVSt.Joost who participated in the experiments in this research.

Last but not least, my friends and family for their constant interest, support, and encouragement.

TiCC Ph.D. Series

- 1 Pashiera Barkhuysen. *Audiovisual Prosody in Interaction*. Promotores: M.G.J. Swerts, E.J. Kraemer. Tilburg, 3 October 2008.
- 2 Ben Torben-Nielsen. *Dendritic Morphology: Function Shapes Structure*. Promotores: H.J. van den Herik, E.O. Postma. Co-promotor: K.P. Tuyls. Tilburg, 3 December 2008.
- 3 Hans Stol. *A Framework for Evidence-based Policy Making Using IT*. Promotor: H.J. van den Herik. Tilburg, 21 January 2009.
- 4 Jeroen Geertzen. *Dialogue Act Recognition and Prediction*. Promotor: H. Bunt. Co-promotor: J.M.B. Terken. Tilburg, 11 February 2009.
- 5 Sander Canisius. *Structured Prediction for Natural Language Processing*. Promotores: A.P.J. van den Bosch, W. Daelemans. Tilburg, 13 February 2009.
- 6 Fritz Reul. *New Architectures in Computer Chess*. Promotor: H.J. van den Herik. Co-promotor: J.W.H.M. Uiterwijk. Tilburg, 17 June 2009.
- 7 Laurens van der Maaten. *Feature Extraction from Visual Data*. Promotores: E.O. Postma, H.J. van den Herik. Co-promotor: A.G. Lange. Tilburg, 23 June 2009 (cum laude).
- 8 Stephan Raaijmakers. *Multinomial Language Learning*. Promotores: W. Daelemans, A.P.J. van den Bosch. Tilburg, 1 December 2009.
- 9 Igor Berezhnoy. *Digital Analysis of Paintings*. Promotores: E.O. Postma, H.J. van den Herik. Tilburg, 7 December 2009.
- 10 Toine Bogers. *Recommender Systems for Social Bookmarking*. Promotor: A.P.J. van den Bosch. Tilburg, 8 December 2009.
- 11 Sander Bakkes. *Rapid Adaptation of Video Game AI*. Promotor: H.J. van den Herik. Co-promotor: P. Spronck. Tilburg, 3 March 2010.
- 12 Maria Mos. *Complex Lexical Items*. Promotor: A.P.J. van den Bosch. Co-promotores: A. Vermeer, A. Backus. Tilburg, 12 May 2010 (in collaboration with the Department of Language and Culture Studies).
- 13 Marieke van Erp. *Accessing Natural History. Discoveries in data cleaning, structuring, and retrieval*. Promotor: A.P.J. van den Bosch. Co-promotor: P.K. Lendvai. Tilburg, 30 June 2010.
- 14 Edwin Commandeur. *Implicit Causality and Implicit Consequentiality in Language Comprehension*. Promotores: L.G.M. Noordman, W. Vonk. Co-promotor: R. Cozijn. Tilburg, 30 June 2010.
- 15 Bart Bogaert. *Cloud Content Contention*. Promotores: H.J. van den Herik, E.O. Postma. Tilburg, 30 March 2011.
- 16 Xiaoyu Mao. *Airport under Control*. Promotores: H.J. van den Herik, E.O. Postma. Co-promotores: N. Roos, A. Salden. Tilburg, 25 May 2011.
- 17 Olga Petukhova. *Multidimensional Dialogue Modelling*. Promotor: H. Bunt. Tilburg, 1 September 2011.
- 18 Lisette Mol. *Language in the Hands*. Promotores: E.J. Kraemer, A.A. Maes, M.G.J. Swerts. Tilburg, 7 November 2011 (cum laude).

- 19 Herman Stehouwer. *Statistical Language Models for Alternative Sequence Selection*. Promotores: A.P.J. van den Bosch, H.J. van den Herik. Co-promotor: M.M. van Zaanen. Tilburg, 7 December 2011.
- 20 Terry Kakeeto-Aelen. *Relationship Marketing for SMEs in Uganda*. Promotores: J. Chr. van Dalen, H.J. van den Herik. Co-promotor: B.A. Van de Walle. Tilburg, 1 February 2012.
- 21 Suleman Shahid. *Fun & Face: Exploring non-verbal expressions of emotion during playful interactions*. Promotores: E.J. Kraemer, M.G.J. Swerts. Tilburg, 25 May 2012.
- 22 Thijs Vis. *Intelligence, Politie en Veiligheidsdienst: Verenigbare Grootheden?* Promotores: T.A. de Roos, H.J. van den Herik, A.C.M. Spapens. Tilburg, 6 June 2012 (in collaboration with the Tilburg School of Law).
- 23 Nancy Pascall. *Engendering Technology Empowering Women*. Promotores: H.J. van den Herik, M. Diocaretz. Tilburg, 19 November 2012.
- 24 Agus Gunawan. *Information Access for SMEs in Indonesia*. Promotor: H.J. van den Herik. Co-promotores: M. Wahdan, B.A. Van de Walle. Tilburg, 19 December 2012.
- 25 Giel van Lankveld. *Quantifying Individual Player Differences*. Promotores: H.J. van den Herik, A.R. Arntz. Co-promotor: P. Spronck. Tilburg, 27 February 2013.
- 26 Sander Wubben. *Text-to-text Generation Using Monolingual Machine Translation*. Promotores: E.J. Kraemer, A.P.J. van den Bosch, H. Bunt. Tilburg, 5 June 2013.
- 27 Jeroen Janssens. *Outlier Selection and One-Class Classification*. Promotores: E.O. Postma, H.J. van den Herik. Tilburg, 11 June 2013.
- 28 Martijn Balsters. *Expression and Perception of Emotions: The Case of Depression, Sadness and Fear*. Promotores: E.J. Kraemer, M.G.J. Swerts, A.J.J.M. Vingerhoets. Tilburg, 25 June 2013.
- 29 Lianne van Weelden. *Metaphor in Good Shape*. Promotor: A.A. Maes. Co-promotor: J. Schilperoord. Tilburg, 28 June 2013.
- 30 Ruud Koolen. *'Need I say More? On Overspecification in Definite Reference.'* Promotores: E.J. Kraemer, M.G.J. Swerts. Tilburg, 20 September 2013.
- 31 J. Douglas Mastin. *Exploring Infant Engagement. Language Socialization and Vocabulary Development: A Study of Rural and Urban Communities in Mozambique*. Promotor: A.A. Maes. Co-promotor: P.A. Vogt. Tilburg, 11 October 2013.
- 32 Philip C. Jackson, Jr. *Toward Human-Level Artificial Intelligence – Representation and Computation of Meaning in Natural Language*. Promotores: H.C. Bunt, W.P.M. Daelemans. Tilburg, 22 April 2014.
- 33 Jorrig Vogels. *Referential choices in language production: The Role of Accessibility*. Promotores: A.A. Maes, E.J. Kraemer. Tilburg, 23 April 2014.
- 34 Peter de Kock. *Anticipating Criminal Behaviour*. Promotores: H.J. van den Herik, J.C. Scholtes. Co-promotor: P. Spronck. Tilburg, 10 September 2014.
- 35 Constantijn Kaland. *Prosodic marking of semantic contrasts: do speakers adapt to addressees?* Promotores: M.G.J. Swerts, E.J. Kraemer. Tilburg, 1 October 2014.

- 36 Jasmina Mari . *Web Communities, Immigration and Social Capital*. Promotor: H.J. van den Herik. Co-promotores: R. Cozijn, M. Spotti. Tilburg, 18 November 2014.
- 37 Pauline Meesters. *Intelligent Blauw*. Promotores: H.J. van den Herik, T.A. de Roos. Tilburg, 1 December 2014.
- 38 Mandy Visser. *Better use your head. How people learn to signal emotions in social contexts*. Promotores: M.G.J. Swerts, E.J. Krahmer. Tilburg, 10 June 2015.
- 39 Sterling Hutchinson. *How symbolic and embodied representations work in concert*. Promotores: M.M. Louwerse, E.O. Postma. Tilburg, 30 June 2015.
- 40 Marieke Hoetjes. *Talking hands. Reference in speech, gesture and sign*. Promotores: E.J. Krahmer, M.G.J. Swerts. Tilburg, 7 October 2015
- 41 Elisabeth Lubinga. *Stop HIV. Start talking? The effects of rhetorical figures in health messages on conversations among South African adolescents*. Promotores: A.A. Maes, C.J.M. Jansen. Tilburg, 16 October 2015.
- 42 Janet Bagorogoza. *Knowledge Management and High Performance. The Uganda Financial Institutions Models for HPO*. Promotores: H.J. van den Herik, B. van der Walle, Tilburg, 24 November 2015.
- 43 Hans Westerbeek. *Visual realism: Exploring effects on memory, language production, comprehension, and preference*. Promotores: A.A. Maes, M.G.J. Swerts. Co-promotor: M.A.A. van Amelsvoort. Tilburg, 10 Februari 2016.
- 44 Matje van de Camp. *A link to the Past: Constructing Historical Social Networks from Unstructured Data*. Promotores: A.P.J. van den Bosch, E.O. Postma. Tilburg, 2 Maart 2016.
- 45 Annemarie Quispel. *Data for all. How professionals and non-professionals in design use and evaluate information visualizations*. Promotor: A.A. Maes. Co-promotor: J. Schilperoord. Tilburg, 15 June 2016.

