Journal of Visual Languages and Computing 25 (2014) 107–116

Contents lists available at ScienceDirect

Journal of Visual Languages and Computing

journal homepage: www.elsevier.com/locate/jvlc

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

Annemarie Quispel [a,b,*], Alfons Maes [b]

[a] AKV\St.Joost Academy of Fine Art and Design, Avans University, Beukenlaan 1, 4834 CR, Breda, Netherlands
[b] Tilburg University, Tilburg Center for Cognition and Communication (TiCC), Warandelaan 2, 5037 AB, Tilburg, Netherlands

ARTICLE INFO

Available online 8 December 2013

Keywords:
Data visualization
Aesthetics
Layout
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.

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.

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 [1]. 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 [2] 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

☆ This paper has been recommended for acceptance by S.-K. Chang.

* Corresponding author.
E-mail addresses: a.quispel@tilburguniversity.edu (A. Quispel), maes@tilburguniversity.edu (A. Maes).

1045-926X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jvlc.2013.11.007

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 [3–5]. As Schön [5] 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 [6]. 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.

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 [2] 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 [7–14].

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 [15]. 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).

1.3. Design choices and esthetic 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 [16–20]. Kosslyn's [21] and Ware's [13,14] 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 [9,22–24]. Also other domains such as education show an interest in the design of visual displays [26]. For an extensive overview of models for effective display design and methods for testing design principles that have been informed by cognitive science, see Ref. [27].

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 esthetic qualities of visual displays and aim to bridge the gap between usability and aesthetics [28–32]. When it comes to the question what esthetic 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 [33], 2D versus 3D [34,35], or having certain characteristics of works of art, like impressionist color palettes or abstract painting-like compositions [36,37]. Cawthon and Vande Moere [25] 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 [38,39], or, on the other extreme, tried to capture esthetic quality in mathematical formulae, such as metrics for characteristics such as symmetry, balance, or complexity [40,41]. All these studies reveal divergent approaches toward the notion of aesthetics. Some treat it as characteristics contributing to clarity (and, implicitly, through clarity to esthetic experience), whereas other models treat aesthetics as design variables contributing directly to attractiveness through some sort of ‘expressiveness’ [42].

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. [43] 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 [11]. 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 [34] 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 [44–46] 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 [47] states that simplicity is highly overrated, and suggests that other factors should be considered as well, such as aesthetics and symbolism. Inbar et al. [48] measured people's preference for standard bar graphs and minimalist versions taken from Tufte [11]. They found that people prefer non-minimalist bar graphs over the minimalist versions, but these were still simple, conventional bar charts. Bateman et al. [33] 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 [49], 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.

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 multi-way 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) [50]. 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 [51] classified graphics as depicting one, two, or three variables. Tufte [11] 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 and McGill [9] and Shneiderman [52] categorized data visualizations according to data type and exploratory task. Other taxonomies, like the one of MacDonald-Ross [2], are functional in nature, and focus on intended use of the diagrams. And yet others developed structural taxonomies, based on the diagrams' physical structure [7,43,53].

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.

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. [43] 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 Fig. A1 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 [7], 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 [54] 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 1 h 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 [11] 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 Fig. A1 (11)) or represented the proportions between parties inaccurately (as in the cup cake example, Fig. A1 (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 Fig. A1 (19)).

2.2. *Expression mode: pictorial vs. abstract*

As Zacks et al. [43] 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. [33] 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’ [55,56]. 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 [57], or iconic vs. symbolic [58]. 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 Fig. A1 (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. Fig. A1 (19)); of the pictorial ones, 8 were inaccurate (e.g. Fig. A1(20)).

3. Evaluation study

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 [34], or minimalist versus non-minimalist but still simple bar and pie charts [48]. 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 so as 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. [33] 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.

3.2. 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).

3.2.1. Participants

Participants were 30 students majoring in graphic design (14 male, 16 female) at AKV|St.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.

3.2.2. Materials

We selected 20 out of the 41 visualizations produced in the production experiment (see Fig. A1). 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.

3.2.3. Procedure

Respondents took part in the experiment individually. The experiment took about 30 min. 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.

3.2.3.1. 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 s. 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.

3.2.3.2. 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.

3.2.3.3. 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 s. Respondents were asked to rate each visualization's perceived clarity. As they performed this task after the information retrieval task, it was supposed that they would base this judgment on the ease with which they had been able to retrieve the information from each visualization.

3.2.3.4. 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.

3.2.3.5. 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.

3.2.3.6. 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 80–100% 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*), and 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. Results

3.3.1. 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.

3.3.1.1. 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 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 ($F<1$).

3.3.1.2. 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 ($F<1$).

3.3.2. Selection task and explanation

3.3.2.1. 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.

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).

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

	Professionals		Laypeople		Overlap	
	n	*Item number*	*n*	*Item number*	*n*	*Item number*
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 3
Reasons mentioned for 'best of' and 'worst of' selection related to design experience (professionals, laypeople). Means per participant (standard deviations between brackets).

	Professionals M (SD)	Laypeople M (SD)
Best		
Clear	.98 (1.19)	2.10 (1.09)
Attractive	2.73 (.63)	2.07 (1.11)
Different	.77 (1.07)	.34 (.66)
Worst		
Unclear	2.13 (1.07)	2.78 (.53)
Unattractive	1.07 (1.05)	.37 (.58)

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 (Fig. A1 (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), do not show any differences in proportions at all.

3.3.2.2. 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 ($t(69)=3.59$, $p=.001$).

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

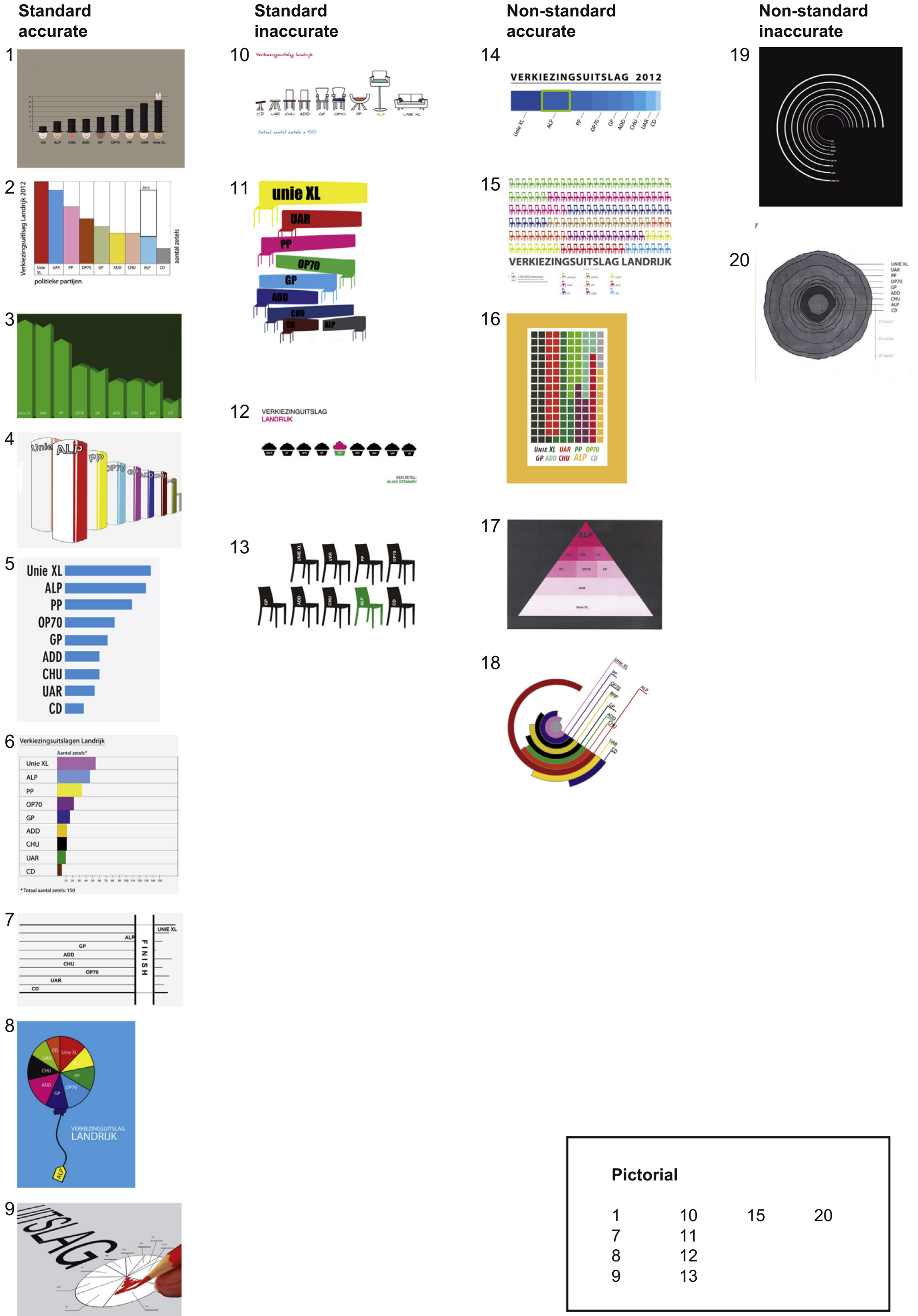

Fig. A1. Sample of 20 visualizations used in the experiment, classified according to type of construction and mode of expression.

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 on the basis of the results found by Bateman et al. [33]. 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 do not 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, disproportionality 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 do not 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 we 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 $\times$ 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.

Appendix

See Appendix Fig. A1.

References

[1] P.K. Storkerson, Communication research: theory, empirical studies, and results, in: A. Bennet (Ed.), Design Studies—Theory and Research in Graphic Design, Princeton Architectural Press, New York, 2006.
[2] M. MacDonald-Ross, How numbers are shown: a view of research on the presentation of quantitative data in texts, AV Commun. Rev. 25 (1977) 359–409.
[3] Michael Polanyi, The Tacit Dimension, Doubleday, Garden City, NY, 1966.
[4] N. Cross, Designerly ways of knowing, Des. Stud. 3 (4) (1982) 221–227.
[5] D.A. Schön, The Reflective Practitioner. How Professionals Think in Action, Templer-Smith, London, 1983.
[6] K. Friedman, Theory construction in design research: criteria, approaches, and methods, Des. Stud. 24 (6) (2003) 507–522.
[7] J. Bertin, Semiology of Graphics: Diagrams Networks Maps, Esri Press, Redlands, CA, 1983.
[8] S.K. Card, J.D. Mackinlay, B. Shneiderman, Readings in Information Visualization. Using Vision to Think, Morgan Kaufmann, San Francisco, CA, 1999.
[9] W.S. Cleveland, R. McGill, Graphical perception: theory, experimentation, and application to the development of graphical methods, J. Am. Stat. Assoc. 79 (1984) 387.
[10] W.S. Cleveland, The Elements of Graphing Data, Wadsworth, Monterey, CA, 1985.
[11] E.R. Tufte, The Visual Display of Quantitative Information, Graphics Press, Cheshire, Connecticut, 1983.
[12] R. Spence, Information Visualization, ACM Press, New York, 2001.
[13] C. Ware, Information Visualization. Perception for Design, Morgan Kaufman, San Francisco, CA, 2004.
[14] C. Ware, Visual Thinking for Design, Morgan Kaufman, Burlington, MA, 2008.
[15] Card, S.K., MacKinlay, J., The structure of the information visualization design space, in: Proceedings of the IEEE Symposium on Information Visualization, 1997.
[16] R. Arnheim, Visual Thinking, University of California Press, Berkely, 1969.
[17] J.D. Mackinlay, Automating the design of graphical presentations of relational information, ACM Trans. Graph. 5 (1986) 110–141.
[18] M. Scaife, Y. Rogers, External cognition: how do graphical representations work? Int. J. Hum.-Comput. Stud. 45 (1996) 115–143.
[19] P.A. Carpenter, P. Shah, A model of the perceptual and conceptual processes in graph comprehension, J. Exp. Psychol.: Appl. 4 (1998) 75–100.
[20] B. Tversky, Visualizing thought, Top. Cognitive Sci. 3 (2011) 499–535.
[21] S.M. Kosslyn, Elements of Graph Design, W.H. Freeman, New York, 1994.

[22] J. Heer, N. Kong, M. Agrawala, 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, ACM, New York, NY, USA, 2009, pp. 1303–1312.
[23] K.B. Lloyd, D.J. Jankowski, A cognitive information processing and information theory approach to diagram clarity: a synthesis and experimental investigation, J. Syst. Software 45 (1999) 203–214.
[24] P. Shah, P. Carpenter, Conceptual limitations in comprehending line graphs, J. Exp. Psychol.: Gen. 124 (1995) 337–370.
[25] N. Cawthon, A. Vande Moere, The effect of aesthetic on the usability of data visualization, in: Proceedings of the IEEE 11th International Conference Information Visualization (IV'07), 2007.
[26] T.Van Gog, K. Scheiter, Eye tracking as a tool to study and enhance multimedia learning, Learn. Instr. 20 (2010) 95–99.
[27] M. Hegarty, The cognitive science of visual-spatial displays: implications for design, Top. Cognitive Sci. 3 (3) (2011) 446–474.
[28] G. Judelman, Aesthetics and inspiration for visualization design: bridging the gap between art and science, in: Proceedings of the International Conference on Information Visualisation (IV), IEEE Computer Society, London, UK, 2004, pp. 245–250.
[29] R. Kosara, Visualization criticism—the missing link between information visualization and art, in: Proceedings of the 11th International Conference on Information Visualisation (IV), IEEE CS Press, Washington, DC, 2007, pp. 631–636.
[30] A. Lau, A. Vande Moere, Towards a model of information aesthetic visualization, in: Proceedings of the IEEE International Conference on Information Visualization (IV'07), 2007, pp. 87–92.
[31] Z. Pousman, J. Stasko, M. Mateas, Casual information visualization: depictions of data in everyday life, IEEE Trans. Vis. Comput. Graph. 13 (6) (2007) 1145–1152.
[32] A. Vande Moere, H. Purchase, On the role of design in information visualization, Inf. Vis. 10 (4) (2011) 356.
[33] S. Bateman, R.L. Mandryk, C. Gutwin, A. Genest, D. McDine, C. Brooks, Useful junk? The effects of visual embellishment on comprehension and memorability of charts, in: Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2010), Atlanta, GA, USA, 2010, pp. 2573–2582.
[34] E. Levy, J. Zacks, B. Tversky, D. Schiano, Gratuitous graphics? Putting preferences in perspective, in: Proceedings of the CHI96, 1996.
[35] N. Tractinsky, J. Meyer, Chartjunk or goldgraph? Effects of presentation objectives and content desirability on information presentation, MIS Q. 23 (1999) 3.
[36] S.I. Fabrikant, S. Rebich-Hespanha, M. Hegarty, Cognitively inspired and perceptually salient graphic displays for efficient inference making, Ann. Assoc. Am. Geogr. 100 (2010) 13–29.
[37] T. Skog, S. Ljungblad, L.E. Holmquist, Between aesthetics and utility: designing ambient information visualizations, in: Proceedings of the InfoVis, 2003, pp. 233–240.
[38] M. Kurosu, K. Kashimura, Apparent usability vs. inherent usability, in: Proceedings of the CHI '95 Conference Companion, 1995, pp. 292–293.
[39] N. Tractinsky, A. Shoval-Katz, D. Ikar, What is beautiful is usable, Interact. Comput. 13 (2) (2000) 127–145.
[40] D.E. Berlyne, Aesthetics and Psychobiology, Appleton-Century-Crofts, New York, 1971.
[41] D. Ngo, L. Teo, J. Byrne, Modelling interface aesthetics, Inf. Sci. 152 (2003) 25.
[42] T. Lavie, N. Tractinsky, Assessing dimensions of perceived visual aesthetics of Web sites, Int. J. Hum.-Comput. Stud. 60 (2004) 269–298.
[43] J. Zacks, E. Levy, B. Tversky, D. Schiano, Graphs in print, in: M. Anderson, B. Meyer, P. Olivier (Eds.), Diagrammatic Representation and Reasoning, Springer Verlag, London, 2002, pp. 187–206.
[44] R. Klanten (Ed.), Data Flow—Visualising Information in Graphic Design, Gestalten, Berlin, 2008.
[45] R. Klanten, S. Ehmann, N. Bourquin, T. Tissot (Eds.), Data Flow 2—Visualising Information in Graphic Design, Gestalten, Berlin, 2010.
[46] D. McCandless, Information is Beautiful, Collins, London, 2009.
[47] D. Norman, Simplicity is highly overrated, Interactions, 14, ACM, New York, NY, 2007.
[48] O. Inbar, N. Tractinsky, J. Meyer, Minimalism in information visualization—attitudes towards maximizing the data-ink ratio, Proceedings of the ECCE, ACM Press, 2007, 185–188.
[49] N. Holmes, Nigel Holmes on Information Design, Jorge Pinto Books, New York, 2006.
[50] J. Zacks, B. Tversky, Bars and lines: a study of graphic communication, Mem. Cognit. 27 (6) (1999) 1073–1079.
[51] W.S. Cleveland, Graphs in scientific publications, Am. Statistician 38 (1984) 4.
[52] B. Shneiderman, The eyes have it: a task by data type taxonomy for information visualization, Presented at Visual Languages 96, 1996, pp. 336–343.
[53] G. Lohse, K. Biolsi, N. Walker, H. Rueler, A classification of visual representations, Commun. ACM 37 (12) (1994) 36–49.
[54] I. Spence, S. Lewandowsky, Displaying proportions and percentages, Appl. Cognitive Psychol. 5 (1991) 61–77.
[55] A. Blackwell, Y. Engelhardt, A meta-taxonomy for diagram research, in: M. Anderson, B. Meyer, P. Olivier (Eds.), Diagrammatic Representation and Reasoning, Springer, London, 2002, pp. 47–64in: M. Anderson, B. Meyer, P. Olivier (Eds.), Diagrammatic Representation and Reasoning, Springer, London, 2002, pp. 47–64.
[56] Y. Engelhardt, The Language of Graphics—A Framework for the Analysis of Syntax and Meaning in Maps, Charts and Diagrams, ILLC, UvA, Amsterdam, 2002.
[57] C.J. Richards, The fundamental design variables of diagramming, in: M. Anderson, B. Meyer, P. Olivier (Eds.), Diagrammatic Representation and Reasoning, Springer, London, 2002, pp. 85–102in: M. Anderson, B. Meyer, P. Olivier (Eds.), Diagrammatic Representation and Reasoning, Springer, London, 2002, pp. 85–102.
[58] B. Tversky A. Kessel, Using gestures and diagrams, in: Proceedings of Sig2 Bi-annual meeting Text and Graphics Comprehension, 2006.