



Evaluating the Efficiency of Physical Visualizations

Yvonne Jansen, Pierre Dragicevic, Jean-Daniel Fekete

► To cite this version:

Yvonne Jansen, Pierre Dragicevic, Jean-Daniel Fekete. Evaluating the Efficiency of Physical Visualizations. Proceedings of the 2013 Annual Conference on Human Factors in Computing Systems (CHI 2013), ACM, Apr 2013, Paris, France. pp.2593-2602, 10.1145/2470654.2481359 . hal-00781831

HAL Id: hal-00781831

<https://hal.inria.fr/hal-00781831>

Submitted on 22 Sep 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Evaluating the Efficiency of Physical Visualizations

Yvonne Jansen

Université Paris-Sud & INRIA

yvonne.jansen@inria.fr

Pierre Dragicevic

INRIA

pierre.dragicevic@inria.fr

Jean-Daniel Fekete

INRIA

jean-daniel.fekete@inria.fr



Figure 1. Examples of physical visualizations: a) electricity consumption over one year with each day split into 30min intervals (Detroit Edison electrical company, 1935); b) data sculpture showing a world map of GDP (wooden base) and derivatives volume (wireframe) (Andreas Nicolas Fischer, 2008); c) 3D bar charts depicting the evolution of country indicators over time built specifically for our study.

ABSTRACT

Data sculptures are an increasingly popular form of physical visualization whose purposes are essentially artistic, communicative or educational. But can physical visualizations help carry out actual information visualization tasks? We present the first infovis study comparing physical to on-screen visualizations. We focus on 3D visualizations, as these are common among physical visualizations but known to be problematic on computers. Taking 3D bar charts as an example, we show that moving visualizations to the physical world can improve users' efficiency at information retrieval tasks. In contrast, augmenting on-screen visualizations with stereoscopic rendering alone or with prop-based manipulation was of limited help. The efficiency of physical visualizations seems to stem from features that are unique to physical objects, such as their ability to be touched and their perfect visual realism. These findings provide empirical motivation for current research on fast fabrication and self-reconfiguring interfaces.

Author Keywords

Physical visualization; evaluation; 3D visualization.

ACM Classification Keywords

H5.2 User Interfaces: Evaluation/Methodology

General Terms

Human Factors; Design; Experimentation; Measurement.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2013, April 27–May 2, 2013, Paris, France.

Copyright 2013 ACM 978-1-4503-1899-0/13/04...\$15.00.

INTRODUCTION

Traditional visualizations map data to pixels or ink, whereas physical visualizations map data to physical form. Physical visualizations recently became popular in the form of *data sculptures*, i.e., data-driven artifacts of various shapes and sizes, ranging from personal jewelry¹ to large museum installations². These are built by artists and designers who seek to elicit emotions and convey meaning beyond mere data [40].

Physical visualizations have also been built for supporting goal-oriented productivity tasks. In the 1930s, two American electricity providers were building physical visualizations to better anticipate power demands (Figure 1a) [34]. In the 1970s, Bertin was building physical adjacency matrix visualizations to study matrix reordering [13, p. 78]. Today, General Motors is using 3D Lego block visualizations to get a better overview of problems in their car production pipeline [48].

These last examples are relevant to the field of information visualization (infovis). In contrast with art and design, infovis is mostly interested in how visualizations can be used to convey objective information about the data itself and yield factual insights about this data. But physical visualizations have generated comparatively little interest in infovis. Apart from a few anecdotal examples they are rarely used by analysts, and they are almost completely ignored in research.

So far this lack of interest could be explained by the remarkable superiority of personal computers over physical matter. Typical computer visualization systems are able to accommodate heterogeneous and dynamic datasets, and support powerful interactive exploration tools like dynamic filtering and search. In contrast, physical visualizations can take time to build and are typically static. However, today this is changing due to two emerging technology trends.

¹e.g., tinyurl.com/weather-bracelet

²e.g., tinyurl.com/mount-fear

The first one is digital fabrication, which makes physical objects increasingly easy to build. Machines such as laser cutters and 3D printers are already being used to create accurate data sculptures [41]. As these machines become faster, there will be more and more situations where building physical visualizations to explore particular datasets will be realistic.

The second technology trend is the increasing amount of work in the field of tangible computing on computationally-augmented and self-reconfigurable interfaces [16, 29]. This suggests that it will eventually become easy to build physical visualizations that can update themselves with new data, and can support interactive exploration tools that are as powerful as the ones we use on our computers today³.

Nevertheless, physical visualizations will only be adopted if they provide clear benefits. While the technological barriers to building them will disappear, it is less clear whether they are really effective for carrying out analytical tasks, and in particular, if they can outperform their virtual counterparts. To our knowledge, this question has never been investigated.

We present the first controlled study comparing physical visualizations to on-screen visualizations from an infovis perspective. We focus on static physical visualizations because if clear benefits are found for these, this will provide a solid motivation for ongoing work on dynamic physical objects. We also focus on 3D visualizations, as these are common in the physical form but problematic to use on computers.

We start with an overview of research in related areas, after which we motivate our experimental design. We then report on a first experiment whose purpose was to verify whether physical visualizations can outperform on-screen visualizations. We then report on a follow-up experiment whose purpose was to understand which unique features of physical visualizations may account for these results. We then conclude with a general discussion and suggestions for future work.

BACKGROUND

We clarify a few concepts then discuss related work, including data sculptures, work from tangible computing, and studies on the manipulation and perception of physical objects.

Physical and On-Screen Modalities

The dichotomy between real/physical and virtual/digital in HCI is an elusive concept [25]. Nevertheless, we roughly define *physical visualizations* as visualizations that are made of physical matter (like in Figure 1), as opposed to presented on a computer screen or projected on a surface as it is traditionally the case. This includes matter whose shape or properties change over time. *Physical modality* and *on-screen modality* will refer to these two types of presentations. Paper & ink is a separate modality that is outside the scope of this study.

Physical Visualizations vs. Models

A *model* is a reproduction of a real object, usually at a different scale. Examples include architectural models and molecule models. In contrast, a *visualization* involves the

process of *visual mapping*, i.e., the choice of visual variables (color, position, size, ...) to encode data [6]. Although both models and visualizations can be displayed on-screen or made physical, our work focuses on visualizations. We consider both pure information visualizations where the data is non-spatial (e.g., social networks or hard drive usage) and scientific visualizations where models and visualizations are combined (e.g., fMRI scans or fluid simulations) [26].

2D vs. 3D Visualizations

When visualizing data, the choice of visual mapping is an issue that is *orthogonal* to the choice of modality. A key part of the visual mapping process consists in mapping data to positions, either in 2D or in 3D. Both 2D and 3D visualizations can be shown on a screen or exist as physical objects. Overall 3D visualizations are not recommended since they are subject to serious occlusion, distortion, and navigation issues [32]. They are however difficult to avoid when the data is inherently 3D, as often in scientific visualization [6]. 3D information visualizations have also been shown to be useful in some tasks [30, 20]. In addition, some of them are widely used and new techniques are regularly being proposed.

However, the controversy of 2D versus 3D only concerns the *on-screen* modality. For the physical modality these issues may be mitigated, as we are used to perceive and manipulate 3D physical objects. In fact, although some physical visualizations are simply extruded 2D visualizations⁴, most of them exploit the three dimensions of space (e.g., Figures 1a & b).

Data Sculptures and Casual Infovis

Data sculptures are by far the most popular and common type of physical visualizations (see, e.g., Figure 1b). There is an extremely large variety of data sculptures, which Vande Moere reviews and discusses extensively [40, 41].

In an article targeted at an infovis audience [40], Vande Moere stresses the limitations of the screen medium and the advantages of information representations that can be “touched, explored, carried, or even possessed”. He argues that data sculptures can convey messages beyond the data itself, can encourage people to reflect on its meaning and change their behavior, and can overall yield a more pleasurable, engaging, and educational experience. He predicts that data communication will progressively be “pushed outside of the digital screen, towards our everyday physical reality experiences”.

While traditional infovis focuses on expert users and task performance, *casual infovis* focuses on large audiences and non-work situations [28]. Vande Moere argues that on-screen visualizations are associated with the former and advocates physical visualizations as an alternative to support the latter [40]. However, on-screen visualizations can also be very effective at supporting casual infovis, for example through the Web medium [28]. Conversely, physical visualizations could also be very effective at supporting work-related activities, but this possibility has so far been little discussed. Notwithstanding, assessing how good people are at reading physical visualizations will inform both traditional and casual infovis.

³Also see tinyurl.com/claytronics for a concept video.

⁴e.g., tinyurl.com/callaghan-vis or tinyurl.com/tohoku-vis

Hybrid Virtual/Physical Visualizations

Although physical visualizations have been mostly ignored in computer science, there has been lots of interest in “hybrid” systems that combine virtual with physical elements. In the fields of tangible computing and scientific visualization, several such systems have been proposed for data exploration.

Physical models and visualizations have been used as props to navigate on-screen scientific visualizations. Hinckley’s neurosurgical planning system uses a physical head doll and a physical plate to control the location of a cutting plane on a virtual model [14]. Kruszynski and Liere [22] propose to use a stylus and 3D printed models of corals to navigate and annotate higher-resolution virtual models of the same corals. Similarly, Konchada et al. [21] propose to use low-resolution 3D printed models and a stylus to help users explore multi-variate volumetric data displayed on a screen. Although these systems suggest that physical imitations of on-screen 3D visualizations may help navigation, these are only used as input devices and are not meant to support visual tasks by themselves. Also, these systems have not been formally evaluated.

Other hybrid systems use co-located virtual and physical elements. Illuminating Clay lets users deform a ductile clay surface and projects data such as land erosion back on the surface [16]. Similarly, URP projects environmental information on physical architectural models [39]. Strata/ICC augments a transparent architectural model of a skyscraper with LEDs visualizing consumption of electricity and water on different floors [38]. These systems present the advantage of supporting both physicality and dynamic data. However, the physical objects involved are models, not visualizations. Abstract data is visualized with projected or embedded light and does not have a physical form. And as before, most of these systems are proofs of concepts and none of them has been evaluated.

Dynamic Physical Visualizations

Some physical visualizations have moving parts that can be manually rearranged, such as Bertin’s adjacency matrices [13] or Lego block visualizations [48]. This adds basic support for “interactivity” but does not adequately support dynamic data. Some actuated data sculptures have been built such as Pulse⁵, a complex live representation of emotions extracted from weblogs, or Wable⁶, a motorized physical bar chart for monitoring RSS feeds. But most of these systems are proofs of concepts that only support very specific datasets.

A promising approach to the physical display of dynamic information are shape displays, often implemented as arrays of actuated bars [18, 23, 27]. Cooperating nanorobots are also being considered [11]. Although these technologies seem to have a high potential for infovis applications, previous work has mostly focused on interaction and hardware issues. Furthermore, there is very little empirical evidence on the effectiveness of these systems compared to visual displays [29]. We address this lack of evidence by assessing the effectiveness of static physical visualizations, expecting that our results will generalize to dynamic physical visualizations.

⁵tinyurl.com/pulse-vis

⁶tinyurl.com/wable

Related Studies

Three types of studies are relevant to physical visualizations: psychology studies on object manipulation and cognition, studies on object manipulation in virtual reality (VR) environments, and studies on 3D rendering and perception.

Physical Objects and Cognition

A vast body of research in educational science and developmental psychology suggests that the manipulation of physical objects can promote understanding and learning [5]. Part of this research is motivated by the *embodied cognition* thesis, according to which cognition is supported by the body and the physical world [47]. However, we do not know of any study comparing physical with on-screen visualizations.

VR Object Manipulation

Virtual reality (VR) is interested in reproducing the physical world, including our natural abilities to manipulate objects. Studies in VR suggest that for 6DOF tasks, 6DOF input devices like physical props are more effective than 2DOF input devices like computer mice [15]. Other factors such as lag or the spatial separation between motor actions and visual output are thought to impact performance [43, 45]. While providing interesting insights, these studies focus on tasks in the motor domain, where vision merely serves object manipulation. They do not easily generalize to tasks in the visual domain, where object manipulation merely serves vision.

3D Rendering and Perception

While the physical world provides a plethora of powerful depth cues [36], computer-generated content needs to create all these cues explicitly. Ware and Franck [44] found that head-coupled stereo viewing and high rendering quality improve path following in large 3D graphs to the point that they outperform 2D graphs. Volumetric displays have been shown to provide the best depth perception but have low resolution [12]. Despite recent advances, current hardware is not yet able to closely mimic the perception of real objects.

Studies Involving Physical Visualizations

McGookin et al. [24] studied how hybrid physical/virtual visualizations and sonification can help visually impaired users read charts. They mention several advantages of the physical modality such as “two-handed interaction, quick overviews, spatial frame of reference and flexibility to employ fingers for marking”. However, their study was only a usability evaluation and they made no comparison with a different modality.

In his thesis, Dwyer [9] compared a physical 3D visualization to printouts of small multiples for infovis tasks, and found that the two differed depending on tasks. Although two different modalities were used, the purpose was to compare visual mappings. Since each modality used a different visual mapping, the study provides no insight on the effect of modality.

Cockburn and McKenzie [7] compared physical versions of DataMountain (2D, 2.5D and 3D) with equivalent on-screen representations. The 3D version performed worse than 2D overall, but the physical modality was found to outperform the on-screen modality. The study however focused on spatial memory and item retrieval times, not on infovis tasks.

STUDY DESIGN RATIONALE

Visualization designs can be compared according to the *visual mapping* they use and/or the *modality* they use, with respect to metrics of interest. The vast majority of infovis studies involve comparisons across visual mappings only. We focus on comparisons across modalities only, a question that to our knowledge has never been investigated. The number of possible comparisons is huge and a single study can only address a small subset. Here we motivate our choices for the datasets, tasks, visual mappings, and interactions used.

datasets

We used country indicator data from Gapminder such as electricity consumptions, birth rates or car mortalities. We generated 16 datasets, each consisting of the value of a country indicator for 10 countries over 10 years (100 values total). These were selected such that: i) country indicators are easy to understand and all different from each other; ii) the variance of values across countries and across years are reasonably high (see Figure 1c); iii) years are evenly spaced.

Choosing small datasets allowed us to keep the experiment complexity under reasonable limits, and to use enough repetitions to level out task difficulty and possible adherence and attachment effects due to specific datasets. We expect that if differences are found for small datasets, they should generalize and maybe even be magnified for larger datasets.

Tasks

Since we are interested in the general usability of on-screen and physical visualizations, we derived our tasks from taxonomies of low-level information retrieval tasks [3, 46]. To keep the length of the experiment manageable we used data gained from pilot studies to converge on 3 different tasks:

1. *Range task*. Indicate the range of values for a given country.
2. *Order task*. Sort the values for a given year ascending.
3. *Compare task*. Locate three given country/year pairs and determine which one has the lowest value.

We were initially interested in including overview tasks such as estimating trends [46] or finding anomalies [3], but they turned out to be hard to operationalize. Nevertheless, our three tasks cover a range of elementary operations such as finding data points or comparing values, and they require some cognitive effort. Each task was expressed in the data domain independently from any visual mapping. For example: “Indicate the range of suicide rates for Denmark”.

We used the 3D on-screen visualization (see below) to devise a *range* task, an *order* task and a *compare* task per dataset. Since the data was real the difficulty of the tasks could not be fully controlled. We chose tasks that were hard enough but feasible when carried out under the 3D on-screen condition, and further conducted a pilot study to level out difficulties.

Visualization

As discussed before we chose to assess the effect of modality when a 3D visualization is used, as we are more likely to find interesting differences than with a 2D visualization. We chose 3D bar charts as they require low “visualization literacy” and are conceptually easy to understand. 3D bar charts

are among the most commonly used 3D visualizations and are supported in most visualization and spreadsheet software. In addition, 3D bar charts can accommodate many different data types such as time series or matrix data. Most important for our purposes, they can be made perceptually similar on both presentation modalities (see Figure 2 b, c) using standard software libraries and digital fabrication technology.

General Visual Design

The 3D bar chart was an array of bars whose heights were proportional to the value of the row/column pair (country and year). We tried several methods for limiting occlusions: a physical chart with transparent acrylic bars, and for the on-screen version, translucent and/or thin bars, showing only the top of the bars, and combinations of both. They all turned out to have poor legibility. We therefore kept the traditional bars.

The countries were ordered using a similarity-based ordering algorithm for adjacency matrices [13]. As color coding is common in 3D bar charts, bars from each country were assigned a separate color using a categorical scale from ColorBrewer2.org. Axis labellings were automatically computed using the algorithm by Talbot et al [35]. In order to facilitate the reading of values and as recommended by Tufte [37, p. 126], tick lines were also displayed on bars.

On-screen Design

The on-screen 3D bar chart visualization (Figure 2b) was developed using Jzy3d, an open source Java/OpenGL library for displaying 3D charts such as bar charts and scatterplots.

The key feature to support with this modality was 3D navigation. Although previous studies suggest that the mouse is not the most adequate device for 3D manipulation and navigation [49, 50, 15], not all 3D tasks require high-DOF control [33]. In our case, tasks do not require to zoom or translate the chart, nor do they require rotations around the axis perpendicular to the screen (roll). Therefore we simply mapped x and y mouse axes to yaw and pitch rotations, a technique that has been shown to be effective for 2-DOF rotations [4].

Although 3D visualization packages typically come with many features, we tried to take these out of the equation as much as possible. This is to facilitate the interpretation of our results, and also because in the future physical models may support similar features. We therefore limited the features to:

- *Label placement*. We use jzy3d’s default text rendering that keeps labels horizontal and oriented towards the viewer. Although the physical modality does not benefit from this feature, it is so common in 3D chart packages that we chose to keep it. In addition, we extended and fine-tuned the placement algorithm so that labels almost never overlap.
- *Scale display*. jzy3d automatically shows and hides scales depending on the chart orientation, so that they never occlude the bars. We chose to keep this feature as well because it is useful and commonly supported in 3D charts.
- *Bar highlighting*. Bars can be marked and unmarked on mouse click. We support this feature because it is common, it greatly facilitates some tasks (especially Compare) and because bars can be touched in the physical modality.

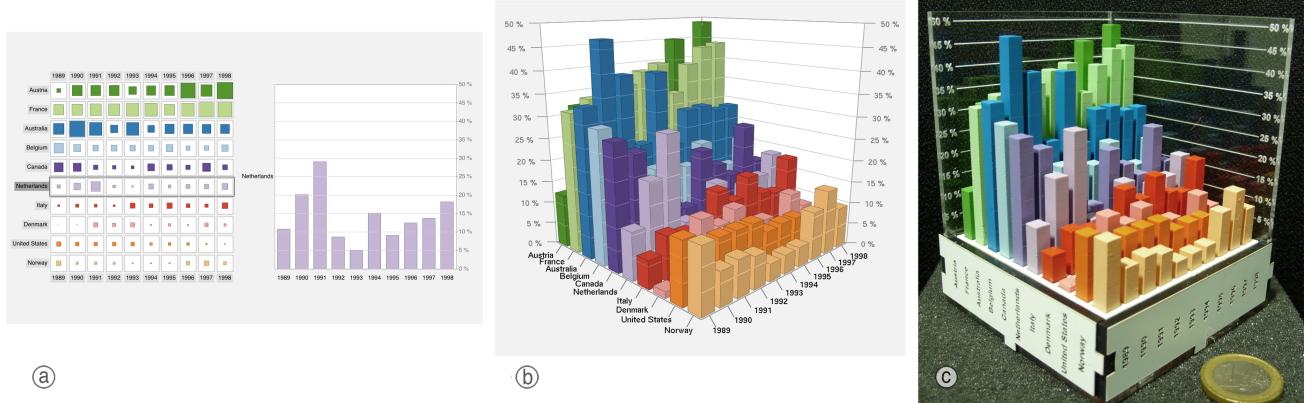


Figure 2. Education expenses data shown under the three conditions: a) on-screen 2D control; b) on-screen 3D bar chart; c) physical 3D bar chart.

- *Perspective switch.* Since it was not clear which projection was best, we let users switch between perspective and orthographic views with the mouse wheel. In the latter mode, the bar chart orientation could snap to the side views.

Considering today’s standard technologies (i.e., desktop computers, current 3D software and passive physical visualizations), we believe these features yield a rather fair comparison between modalities. Removing those features would have biased the comparison towards the physical modality, since today they are standard on computers but hard or impossible to support on physical visualizations.

Physical Design

The physical 3D bar charts (Figure 2c) were replicates of the on-screen 3D bar charts in terms of layout, colors and proportions, with a few minor changes such as label placement. They were built so that they could be held and turned around in a similar way to their on-screen counterparts.

The bar charts were made of laser-cut acrylic. Laser stencils were automatically generated from the data to ensure accurate visualizations. For each country in a given dataset, a 2D bar chart slice was cut ($5 \times 5\text{mm}$ for each bar, 2mm spacing between bars). Each of these slices was then spray-painted. Finally, Tufte’s bar lines [37] were engraved on two sides.

In addition, for each 3D bar chart two scales were made from transparent acrylic sheets, on which axis labels and lines were engraved. The base of the model was built from five pieces of acrylic. Country and year labels were engraved on all four sides of the base, using a vertical orientation. Although this may require rotating the object slightly to facilitate reading, it makes the object more compact and easier to handle.

All pieces were assembled then glued together to make the object feel more sturdy. The outer dimensions were $8 \times 8\text{cm}$, with a weight ranging from 270g to 350g depending on the dataset. A total of 13 such charts were made for the study.

Additional Control Conditions

Although the focus of our study was to compare between the on-screen and the physical setups described above, we added two control conditions to use as baselines of comparison: a 3D stereoscopic condition and an interactive 2D condition.

Stereoscopic Control Condition

On-screen 3D charts provide depth cues through perspective and structure from motion. To assess the benefits of extra depth cues, we added a condition with stereoscopic rendering.

We used quad-buffered stereo rendering in OpenGL and presented the stereo images on a HP 2311 gt 23”, a 3D monitor based on polarized horizontal interlacing that only requires passive glasses. We removed the orthographic mode, because it is not possible to provide stereoscopic cues that are consistent with an infinite viewing distance. Since the HP display can also be used as a normal monitor, the same display was used for the normal (mono) on-screen condition.

2D Control Condition

Although our goal is *not* to compare 2D with 3D visualizations, we included an on-screen interactive 2D condition as a comparison baseline. We tested three such designs:

1. *Line charts.* Superimposed line charts are a common technique for displaying our type of data, but consistent with previous findings [17], ten different overlapping time series caused too much visual clutter.
2. *Small multiples.* Cutting the 3D bar chart across its two main axes produces twenty 2D bar charts. We displayed all of them, together with a magnified view. This option turned out to be confusing because the data was duplicated and it was hard to mentally switch between the two axes.
3. *Matrix.* The whole dataset can be displayed as a matrix and the values shown with squares within cells [10]. Since it is difficult to compare squares precisely, we let users select columns or rows to get the corresponding 2D bar chart.

We chose the last approach as it seemed to be the most effective (Figure 2a). It is also consistent with the 3D chart in several respects: i) it supports the same types of datasets; ii) it has similar scalability; iii) it is conceptually similar: the matrix view is analogous to a top view of the 3D model and the 2D bar chart view is analogous to a side cut.

Users could click and cross [1] columns and row labels to update the 2D bar chart view. As in the 3D chart, they could also highlight individual bars, either on the matrix or on the bar chart view. The axis labeling, chart proportions, Tufte’s lines, colors and visual footprint were similar to the 3D chart.

FIRST EXPERIMENT

The goal of this first experiment was to assess the efficiency of physical 3D visualizations with respect to equivalent 3D on-screen visualizations for information retrieval tasks.

Techniques

In our study we define a *technique* as a combination of a *visual mapping* and a *modality*, with all associated interaction techniques. We included four such techniques (see Table 1, and the previous section for details).

Technique	Visual Mapping	Modality	Illustration
<i>physical</i>	3D bar chart	physical	Figure 2c
<i>stereo</i>	3D bar chart	on-screen stereo	
<i>mono</i>	3D bar chart	on-screen	Figure 2b
<i>2D</i>	matrix + 2D bar chart	on-screen	Figure 2a

Table 1. The four technique conditions.

Tasks

Tasks consisted in answering questions about country indicators, as explained in the previous section. We used 3 types of questions corresponding to the *range* task, the *order* task and the *compare* task described previously.

Procedure

Subjects were first given initial instructions and explained the three types of questions. They were then tested for correct stereovision using a subset of Julesz' random-dot test [19].

Subjects were then presented the 4 techniques one after the other. With every change of technique, they were explained the technique and performed a training run on a dataset different from the experimental datasets, where they practiced answering each type of question. They then saw 3 datasets in sequence. With each change of dataset, they were briefly explained the country indicator and its meaning. They then had to answer 3 questions, one per task type.

Below the question a “Start” button was displayed, and pressing it displayed the possible answers. For the *range* task, subjects had to set two sliders labeled like the axes. For the *order* question, they had to press 10 buttons labeled with country names in the right order. For the *compare* question, they had to press the correct button among three. These buttons were labeled with the values to search for (e.g., “Spain in 2010”) and were revealed only after “Start” was pressed.

Subjects were initially instructed to read the questions carefully before hitting “Start” and turning to the visualization. They were then asked to be as fast and accurate as possible before pressing “Done”. Each time a question was displayed, a message was displayed below the question to remind them of these instructions.

All instructions, questions, and possible answers were given on a separate touch tablet. All subjects agreed to be videotaped. The experiment lasted one hour on average.

Experimental Setup

Figure 3 shows the experimental setup for all *technique* conditions. Physical bar charts were placed on a foam block by the experimenter. The computer mouse and foam block were repositionable. The computer mouse, 3D glasses and physical charts were present only during the relevant conditions.

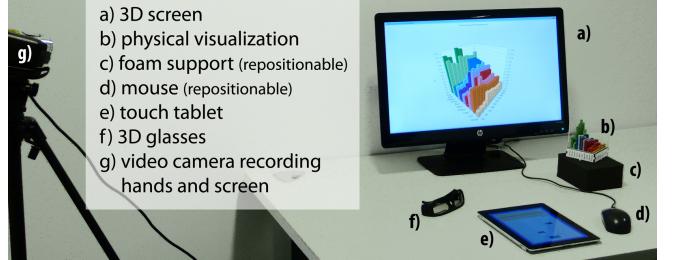


Figure 3. Setup for the first experiment.

Measures

We used two measures of performance: *time on task* and *error rate*. The *time on task* was the interval between the press on “Start” and the press on “Done”. All errors were normalized between 0 and 1. For the *range* task, the error was the average absolute difference between the entered min & max values to the true values, divided by the total axis range. For the *order* task, the error was the normalized Kendall Tau distance (i.e., the number of pairwise disagreements) between the answer and the true order. For the *compare* task, the error was 0 or 1, depending on whether the answer was correct or not.

Since all tasks were feasible with a low error rate under all conditions, time and errors should be linked by a speed-accuracy trade-off and should both capture task difficulty equally well. We instructed subjects to be accurate, and therefore we do not expect to find sizable differences in terms of errors. If this is true this will allow us to base our analysis on time, a more sensitive measure than error rate.

Participants

16 subjects (7 female, mean age 29) were recruited from our university campus. We considered that experience in solving infovis tasks might influence subjects performance and therefore recruited half of our participants from researchers in the field of infovis. All had perfect or corrected to perfect vision and successfully completed our stereo-vision test.

Design

The mapping between datasets and techniques was counterbalanced across subjects by keeping the order of datasets constant and having the presentation order of the techniques follow a balanced latin square (yielding 4 subject groups).

The main factor was *technique* (physical, stereo, mono, 2D). Secondary factors were *infovisbg* (yes, no), *group* (1..4) and *task* (range, order and compare). Therefore we had 16 participants \times 4 techniques \times 3 datasets \times 3 tasks = 576 questions with performance measures.

Hypotheses

Prior to the experiment we ran a pilot with four subjects to check the clarity of the instructions and to get initial estimates of effect sizes. Our hypotheses are based on this pilot:

H1 Task time with *physical* is about 15–20% lower than with both *mono* and *stereo*.

H2 *2D* outperforms all other techniques by no more than 50% in time.

H3 *stereo* is slightly faster than *mono*.

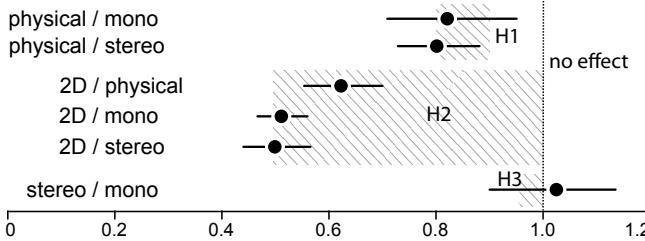


Figure 4. Time ratios between techniques, with 95% CIs. Hatched areas indicate expected effect sizes as expressed in our hypotheses.

Results

A repeated measures ANOVA for *error rate* only revealed an effect of *task* ($F_{(2,24)} = 21.44$, $p < 0.001$). Since *technique* exhibited no measurable effect and errors were generally low ($M_r = 0.017$, $M_o = 0.018$, $M_c = 0.141$), we focus our analysis on time. All *time on task* measures were log-transformed to correct for skewness [31]. All reported means are antilogged and therefore indicate geometric means [31]. Reported differences in pairwise comparisons are also antilogged and therefore indicate *ratios* between geometric means.

A repeated measures ANOVA for *time on task* revealed an effect of *technique* ($F_{(3,24)} = 165.18$, $p < 0.001$). Figure 4 summarizes the differences together with our initial hypotheses. The x-axis shows time ratios between techniques (times being geometric means of all *time on task* measures), e.g., 0.8 means the first technique takes on average 20% less time than the second. Intervals indicate all plausible values, their midpoint being about 7 times more likely than their endpoints [8].

Overall our data is consistent with our hypotheses, with a few uncertainties as to the actual effect sizes. *Physical* may be slightly faster than expected. *2D* may be slightly faster too, but not compared to *physical*. Contrary to *H3*, there is no evidence of *stereo* outperforming *mono*. Overall, we can safely conclude that the physical 3D visualization is more efficient than its 3D on-screen counterpart, *mono* or *stereo* alike.

Task also had a significant effect on *time on task* ($F_{(2,16)} = 149.06$, $p < 0.001$). As Figure 5 shows, the effect was consistent across tasks with an interaction between *technique* and *task*, i.e., the advantage for the *2D* technique was less pronounced for *compare*. So although *2D* beats all 3D conditions, the effect is weaker if a task cannot be solved by one *2D* cut. Neither *infovisbg* (infovis background) nor *group* (ordering of techniques) had a significant effect on *time on task*.

User Feedback

When asked to rank techniques according to preference, subjects gave median rankings that were consistent with technique speed. One subject mentioned finding the *2D* matrix very efficient but preferred the physical chart in terms of “fun and comfort”. Another found the physical chart “very easy to control” and “easier to focus on the desired part of the chart” compared to on-screen 3D, where “I would loose my ‘mark’ easily during rotation”. One subject noted that for dealing with occlusion in 3D, a cutting tool similar to the *2D* condition would be useful (including for the physical chart, with the use of “LEDs and sensors”). One subject found it hard to read labels in *stereo* while another reported feeling “dizzy”.

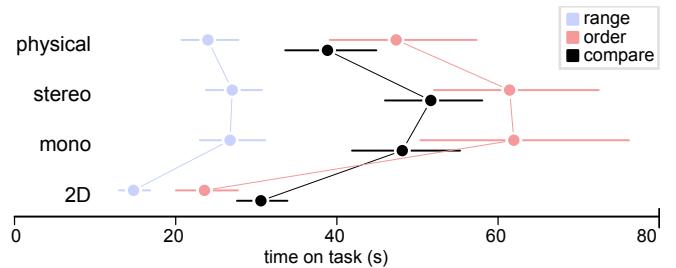


Figure 5. Average time per technique and task, with 95% CIs.

Observations from Video recordings

The video recordings revealed differences between subjects in how they approached the physical bar charts. Many were hesitant at first and inspected them while leaving them on the foam support. Eventually, they picked them up which visibly increased comfort, especially for reading vertical labels. Most subjects used their fingers to temporarily “mark” rows and columns of interest. Eventually, almost everyone switched to putting their fingers directly on the bars relevant to the task. It seemed that the sooner subjects converged on this strategy the faster they finished the tasks.

Discussion

Our data clearly shows that physical 3D bar charts outperform on-screen 3D bar charts. While there can be many contributing factors, three of them seem of interest:

The Role of Touch

Touch seemed to play a major role with the physical chart, as fingers were used to “mark” parts of the chart that were relevant to the task, and therefore seemed to serve as external cognitive and visual aids. We did not expect such high a frequency and variety of uses. Examples included using fingers to relocate previously identified items (for *compare*), following paths (finding country/year intersections in *compare*), focusing on subsets of interest (year in *range* and countries in *order*) and maintaining states (sorted/unsorted bars in *order*).

On the on-screen bar chart, similar marking actions could be emulated by highlighting bars but those actions had to be sequential: users had to stop rotating the bar to perform marking actions, and bars needed to be marked one by one. Occluded bars were also impossible to mark while on the physical chart fingers could reach behind bars. Hence it is plausible that the physical modality was faster because marking actions were essential, and because these actions could be carried out more efficiently, more flexibly, with less attention and concurrently with other actions. Although fingers could have occasionally occluded the chart, proprioceptive information may have compensated for this (feeling bar heights in *compare*).

Direct vs. Indirect Rotations

Our tasks required rotating the chart for performing visual search (e.g., countries or year labels in all tasks), 3D visual comparison (e.g., bar heights in *compare*), and dealing with occlusions (in *order* and *compare*). 2DOF rotations were appropriate for the on-screen modality and we used the best known mouse technique [33, 4]. Charts could be rotated smoothly and rapidly without clutching, which was more difficult with the physical chart, even with two hands.

Despite the ease of use of the mouse, its indirect mapping could have made it harder to use. Indeed, one subject felt it was more difficult to visually track items under this condition. There is evidence that direct physical rotation is more efficient than mouse rotation [15] but results only apply to 3DOF rotation tasks in the motor domain (rotating to specific viewpoints). Nevertheless, it remains possible that the physical modality was faster partly because direct rotation provided better support for some tasks than 2DOF mouse rotation.

Visual Realism

Our tasks being in the visual domain, vision must play a major role. The two modalities follow the same visualization design but differ in several respects: i) resolution (1920×1080 vs. ∞ but a precision of about 0.5mm); ii) stereoscopic cues (except for *stereo*); iii) accommodation cues; iv) shading and shadows; v) texture (none vs. spray painting imperfections). Regarding resolution, all tasks have been designed so it is sufficient in both modalities. Regarding stereoscopic cues, our experiment suggests they are of limited help *per se*. However, since the physical chart is visually more realistic in many respects, this could have made visual tasks more comfortable.

Questions

We identified three factors that could possibly account for the superiority of the physical modality. We therefore chose to address the following questions in a second experiment:

1. *How important is touch in the physical modality?* To answer this we will compare the previous *physical* condition with a condition where subjects are instructed not to touch.
2. *What is the relative importance of direct rotations?* For this we will compare the previous *mono* condition with a condition that employs prop-based input.
3. *What is the relative importance of visual realism?* For this we will compare the no-touch physical condition above with the on-screen prop-based condition.

SECOND EXPERIMENT

The goal of the second experiment was to further investigate factors contributing to the efficiency of physical visualizations. For this we added an “enhanced” version of the on-screen and an “impoverished” version of the physical chart.

Techniques

We used the four following techniques:

- *touch*: same as *physical* from the first experiment, except touch was explicitly encouraged in the instructions.
- *no touch*: same as *physical* except subjects were told not to use their fingers to mark points of interest (labels & bars).
- *prop*: same as *mono* from the first experiment but with tangible prop control, and without bar highlighting.
- *mouse*: same as *mono* but with bar highlighting disabled.

The following chain illustrates our planned comparisons and the corresponding effects we are interested in.

$$\text{mouse} \xrightarrow{+ \text{direct rotation}} \text{prop} \xrightarrow{+ \text{visual realism}} \text{no touch} \xrightarrow{+ \text{touch}} \text{touch}$$

In *mouse* and *prop* we deactivated bar highlighting to avoid confounds. No stereoscopic rendering was used as the first experiment failed to provide clear evidence for advantages.

The physical prop was a regular physical chart marked “prop” with a tracker attached underneath. Subjects were told to ignore its data and refer to the screen. The tracker was housed in an acrylic case containing a Sparkfun Razor IMU 3DOF rotation sensor, an XBee board and a LIPO battery, adding $8 \times 8 \times 1\text{cm}$ in size and 40g in weight. The sensor communicated wirelessly with the PC showing the chart. The total system lag, estimated by video frame counting, was 100-150ms. The sensor was calibrated such that the on-screen chart mirrored the physical chart as seen by the user. The same sensor was attached to all physical bar charts to avoid confounds.

Modifications to the Experimental Design

To get more sensitive measures we only used the most cognitively demanding task, i.e., *compare*. New questions were chosen such that there was no viewpoint with all three bars visible at the same time. The values were also too close together to reliably determine the lowest bar from a top view.

The procedure and experimental setup were similar to the previous experiment. In the *prop*, *touch* and *no touch* conditions, the rotation sensor was recalibrated before each new data set. In the *touch* and *no touch* conditions, a reminder of whether touch was allowed or not was displayed before each question.

We recruited 16 subjects (2 female, mean age 29), 8 of which were randomly chosen from our previous pool. We did not control for infovis experience as we previously did not find any effect, and instead compared former with new subjects.

Our main factor was again *technique*. Secondary factors were *former subject* and *group*. We had $16 \text{ participants} \times 4 \text{ techniques} \times 2 \text{ datasets} \times 4 \text{ questions} = 512$ questions with performance measures. We again measured time and errors.

Hypotheses

Effect sizes are based on a pilot study involving 4 participants.

H1 *touch* requires 15-25% less time than *no touch*.

H2 *no touch* requires at least 10% less time than *prop*.

H3 *prop* and *mouse* differ by no more than 5%.

Results

Error rates were low ($M=0.16$) and a repeated measures ANOVA showed no effect of *technique*, so we again focus on time. A repeated measures ANOVA for *time on task* showed no statistically significant effect for *former subject* or *group*, but *technique* had a clear effect ($F_{(3,24)} = 36.04$, $p < 0.001$).

Figure 6 on the next page shows the results of our planned comparisons together with our hypotheses. Overall our data is consistent with our hypotheses but effect sizes may be slightly smaller than we expected. The possibility of *prop* being slightly faster than *mouse* should also be considered. Figure 7 on the next page shows mean times per technique.

Discussion

Our follow-up experiment suggests that *direct rotation* is not critical for information retrieval on 3D visualizations. The gains are at best modest (3% less time CI [-4%, 10%]). This seems inconsistent with previous VR studies [15], but recall these only consider 3DOF rotation tasks. For visual inspection tasks, 2DOF mouse control seems appropriate [4].

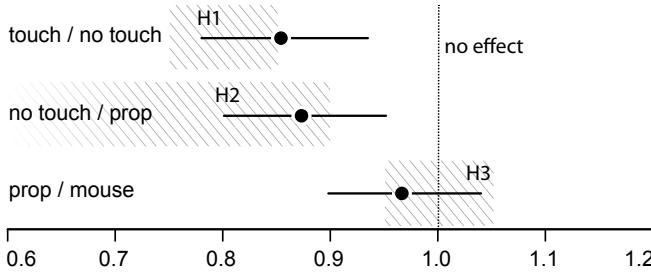


Figure 6. Time ratios between techniques, with 95% CIs. Hatched areas indicate expected effect sizes as expressed in our hypotheses.

It is unclear whether the prop condition could have been dramatically improved. The *total* lag was 100-150ms, but these values are low in VR standards and probably have little influence [2]. Filters were fine-tuned so that the chart was responsive when rotated and stood still during visual examination. Other form factors could have been considered [50] but form factor might not be critical [45]. Elaborate mirror setups are possible that can co-locate motor input and visual output. Evidence shows they do not help translation tasks but could facilitate rotation tasks [45]. However, improvements are less clear when using an egocentric calibration as we did [42].

Overall, it seems that our prop emulated the physical chart reasonably well in terms of rotation control. It was imperfect and likely slightly less efficient for demanding motor tasks, but these motor tasks are very different from our visual information retrieval tasks. This together casts doubts on the ideas that i) the benefits provided by our physical 3D bar charts mostly lie in their facility of being manipulated, and that ii) these benefits can be brought to screen setups using props.

Our experiment confirms that an advantage of the physical bar chart lies in its ability of being *touched*. We do not refer to the use of *tactile* sensory information but to the *action* of placing fingers to use them as visual or memory aids. These actions effectively allowed subjects to unload cognitive effort into the physical world [47]. We found clear improvements when physical charts could be touched (15% less time CI [7%, 22%]). Participants later reported they felt much more comfortable being able to use their fingers. Some even found it “frustrating” when they were not allowed to touch.

Mouse selection only poorly emulates finger marking. Improvements are possible, but reproducing all sensory cues provided while touching a real object is hard. On a multi-touch screen, fingers would not move with the chart. Using a touch-sensitive prop to select bars can be hard if fingers are not visible on the screen. Such feedback could be added using video inlays or mirrors, but these setups are complex, and correct proprioceptive feedback would still be missing without the use of haptic technology or data-accurate props.

That the physical non-touch condition outperformed the prop condition was also insightful (13% less time CI [5%, 20%]). Although the prop condition was an imperfect emulation of physical manipulation, it is unlikely that this alone can explain the difference. The two conditions also differed in their degree of *visual realism*. We previously did not find clear benefits of stereoscopic cues alone, but multiple depth cues may still facilitate information retrieval in 3D visualisations.

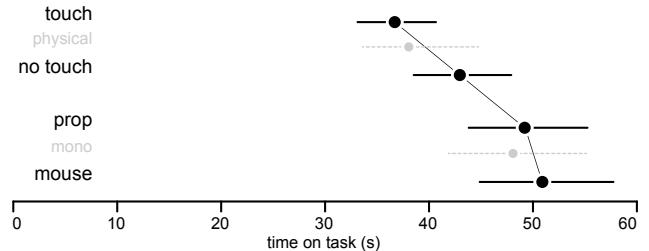


Figure 7. Mean times per technique, with 95% CIs. Results from our first experiment (*task compare*) have been included for reference (gray).

LIMITATIONS OF THE STUDY

Visual Mappings. Our study focused on bar charts. While we expect our findings to generalize to other 3D visualizations, more visual mappings need to be tested across modalities.

Modalities. We only partially investigated the continuum between the physical and on-screen modalities, and more research is needed to understand how they differ. Other modalities also need to be studied. Since touch seems to be an essential cognitive and visual aid, it is likely that 2D visualizations would benefit from paper-based or touchscreen setups.

Tasks. The low-level tasks we used are not targeted at 3D visualizations, although 2D seems less beneficial if tasks cannot be solved with a single 2D cut (Figure 4). We expect this effect to increase for higher-level overview tasks, but further studies are necessary. Also, cost-benefit analyses involving factors other than pure performance [40] are needed to assess when physical 3D visualizations are most useful in practice.

Implications for Design. Our findings suggest that i) if an analyst wishes to use a 3D visualization, making it physical is an option worth considering; ii) physical visualizations should be built to support direct touch and not enclose data⁷. However, the design of effective physical visualizations (size, materials, etc.) is an issue outside the scope of this paper.

CONCLUSION

We presented the first study on the efficiency of physical visualizations. We showed that physical 3D bar charts outperform their on-screen counterparts for information retrieval. Physical touch seems to be an essential cognitive aid, while being able to physically manipulate charts seems comparatively less important. Visual realism might also play a role. All these features seem hard to faithfully reproduce in a virtual setup.

Our results suggest that even passive physical visualizations can be useful and building them with digital fabrication tools seems appropriate, both for research and for personal use. We believe that research on shape-changing surfaces and materials will eventually allow to combine the power of computing with the unique features of physical visualizations.

ACKNOWLEDGMENTS

We thank Basak Alper, Benjamin Bach, Nadia Boukhelifa, Olivier Chapuis, Fanny Chevalier, Geoff Cumming, Tobias Isenberg, Evelyne Lutton, Charles Perin, Andre Spritzer and Barbara Tversky for their helpful feedback on this paper, and Martin Pernollet for his help with Jzy3d.

⁷See, e.g., subsurface crystal engravings at tinyurl.com/crystalvis.

ADDITIONAL MATERIAL

The datasets, experimental data, and a video illustrating the experimental setups, tasks, as well as excerpts of participant video logs are available at www.aviz.fr/phys.

REFERENCES

1. Accot, J., and Zhai, S. More than dotting the i's — foundations for crossing-based interfaces. In *CHI'02, CHI '02* (2002), 73–80.
2. Adamoli, A., Jovic, M., and Hauswirth, M. Lagalyzer: A latency profile analysis and visualization tool. In *Performance Analysis of Systems & Software (ISPASS), 2010 IEEE International Symposium on*, IEEE (2010), 13–22.
3. Amar, R., Eagan, J., and Stasko, J. Low-level components of analytic activity in information visualization. In *InfoVis'05* (2005).
4. Bade, R., Ritter, F., and Preim, B. Usability comparison of mouse-based interaction techniques for predictable 3d rotation. In *Smart Graphics*, Springer (2005), 924–924.
5. Bara, F., Gentaz, E., Colé, P., and Sprenger-Charolles, L. The visuo-haptic and haptic exploration of letters increases the kindergarten-children's understanding of the alphabetic principle. *Cognitive development* 19, 3 (2004), 433–449.
6. Chi, E. H. A taxonomy of visualization techniques using the data state reference model. In *INFOVIS'00* (2000).
7. Cockburn, A., and McKenzie, B. Evaluating the effectiveness of spatial memory in 2d and 3d physical and virtual environments. In *CHI'02* (2002).
8. Cumming, G. *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. Multivariate Applications Series. Routledge, 2011.
9. Dwyer, T. *Two-and-a-half-dimensional Visualisation of Relational Networks*. PhD thesis, University of Sydney, 2004.
10. Elmqvist, N., Do, T.-N., Goodell, H., Henry, N., and Fekete, J.-D. Zame: Interactive large-scale graph visualization. In *Proceedings of the IEEE Pacific Visualization Symposium* (2008), 215–222.
11. Goldstein, S. C., Campbell, J. D., and Mowry, T. C. Programmable matter. *IEEE Computer* 38, 6 (June 2005), 99–101.
12. Grossman, T., and Balakrishnan, R. An evaluation of depth perception on volumetric displays. In *AVI'06* (2006).
13. Henry, N. *Exploring large social networks with matrix-based representations*. PhD thesis, Université Paris-Sud and University of Sydney, 2008.
14. Hinckley, K., Pausch, R., Goble, J. C., and Kassell, N. F. Passive real-world interface props for neurosurgical visualization. In *CHI'94* (1994).
15. Hinckley, K., Tullio, J., Pausch, R., Proffitt, D., and Kassell, N. Usability analysis of 3d rotation techniques. In *UIST'97* (1997).
16. Ishii, H., Lakatos, D., Bonanni, L., and Labrune, J.-B. Radical atoms: beyond tangible bits, toward transformable materials. *interactions* 19, 1 (Jan. 2012), 38–51.
17. Javed, W., McDonnel, B., and Elmqvist, N. Graphical perception of multiple time series. *Visualization and Computer Graphics, IEEE Transactions on* 16, 6 (2010), 927–934.
18. Jonpasang. Hypermatrix. <http://vimeo.com/46857169>, 2012.
19. Julesz, B. *Foundations of cyclopean perception*. U. Chicago Press, 1971.
20. Kjellin, A., Pettersson, L. W., Seipel, S., and Lind, M. Evaluating 2d and 3d visualizations of spatiotemporal information. *ACM Trans. Appl. Percept.* 7, 3 (June 2008), 19:1–19:23.
21. Konchada, V., Jackson, B., Le, T., Borazjani, I., Sotiropoulos, F., and Keefe, D. F. Supporting internal visualization of biomedical datasets via 3d rapid prototypes and sketch-based gestures. In *I3D '11* (2011).
22. Kruszynski, K. J., and Liere, R. V. Tangible props for scientific visualization: concept, requirements, application. *Virtual Reality* 13 (2009), 235–244.
23. Leithinger, D., Lakatos, D., DeVincenzi, A., Blackshaw, M., and Ishii, H. Direct and gestural interaction with relief: A 2.5 d shape display. In *UIST'11* (2011).
24. McGookin, D., Robertson, E., and Brewster, S. Clutching at straws: using tangible interaction to provide non-visual access to graphs. In *Proceedings of the 28th international conference on Human factors in computing systems*, ACM (2010), 1715–1724.
25. Milgram, P., and Kishino, F. A taxonomy of mixed reality visual displays. *IEICE Trans. Information Systems* E77-D, 12 (1994).
26. Munzner, T. Information visualization. Springer-Verlag, Berlin, Heidelberg, 2008, ch. Process and Pitfalls in Writing Information Visualization Research Papers, 134–153.
27. Poupyrev, I., Nashida, T., Maruyama, S., Rekimoto, J., and Yamaji, Y. Lumen: interactive visual and shape display for calm computing. In *SIGGRAPH'04 Emerging technologies* (2004).
28. Pousman, Z., Stasko, J., and Mateas, M. Casual information visualization: Depictions of data in everyday life. *TVCG* 13, 6 (2007).
29. Rasmussen, M. K., Pedersen, E. W., Petersen, M. G., and Hornbæk, K. Shape-changing interfaces: a review of the design space and open research questions. In *CHI'12* (2012).
30. Risden, K., Czerwinski, M. P., Munzner, T., and Cook, D. B. An initial examination of ease of use for 2d and 3d information visualizations of web content. *International Journal of Human-Computer Studies* 53, 5 (2000), 695 – 714.
31. Sauro, J., and Lewis, J. R. Average task times in usability tests: what to report? In *CHI'10* (2010).
32. Shneiderman, B. Why not make interfaces better than 3d reality? *IEEE Comput. Graph. Appl.* 23, 6 (Nov. 2003).
33. Smith, G., Stuerzlinger, W., Salzman, T., Watson, B., and Buchanan, J. 3d scene manipulation with 2d devices and constraints. In *Graphics Interface* (2001), 135–142.
34. Spielman, H. A. “Virtual Reality” Circa 1935. <http://tinyurl.com/spielman2006>, October 2006.
35. Talbot, J., Lin, S., and Hannahan, P. An extension of wilkinson’s algorithm for positioning tick labels on axes. *INFOVIS'10* (2010).
36. Todd, J. The visual perception of three-dimensional structure from motion. *Handbook of perception and cognition, Volume 5: Perception of space and motion* (1995), 201–226.
37. Tufte, E. R. *The visual display of quantitative information*. Graphics Press, Cheshire, CT, USA, 1986.
38. Ullmer, B., Kim, E., Kilian, A., Gray, S., and Ishii, H. Strata/icc: physical models as computational interfaces. In *CHI EA '01* (2001).
39. Underkoffler, J., and Ishii, H. Urp: a luminous-tangible workbench for urban planning and design. In *CHI'99* (1999).
40. Vande Moere, A. Beyond the tyranny of the pixel: Exploring the physicality of information visualization. In *IV'08* (2008).
41. Vande Moere, A., and Patel, S. Analyzing the design approaches of physical data sculptures in a design education context. In *Visual Information Communications International (VINCI'09)* (2009).
42. Ware, C., and Arsenault, R. Frames of reference in virtual object rotation. In *APGV '04* (2004).
43. Ware, C., and Balakrishnan, R. Reaching for objects in vr displays: lag and frame rate. *ACM Trans. Comput.-Hum. Interact.* 1, 4 (Dec. 1994), 331–356.
44. Ware, C., and Mitchell, P. Reevaluating stereo and motion cues for visualizing graphs in three dimensions. In *APGV '05* (2005).
45. Ware, C., and Rose, J. Rotating virtual objects with real handles. *ACM Trans. Comput.-Hum. Interact.* 6, 2 (June 1999), 162–180.
46. Wehrend, S., and Lewis, C. A problem-oriented classification of visualization techniques. In *Visualization'90* (1990).
47. Wilson, M. Six views of embodied cognition. *Psychonomic Bulletin & Review* 9 (2002), 625–636.
48. Wilson, M. How GM is saving cash using legos as a data viz tool. <http://tinyurl.com/mwilson2012>, April 2012.
49. Zhai, S., and Milgram, P. Quantifying coordination in multiple dof movement and its application to evaluating 6 dof input devices. In *CHI'98* (1998).
50. Zhai, S., Milgram, P., and Buxton, W. The influence of muscle groups on performance of multiple degree-of-freedom input. In *CHI'96* (1996).