If Your Mind Can Grasp It, Your Hands Will Help

Simon Stusak
University of Munich (LMU)
Munich, Germany
simon.stusak@ifi.lmu.de

Moritz Hobe
University of Munich (LMU)
Munich, Germany
hobem@cip.ifi.lmu.de

Andreas Butz
University of Munich (LMU)
Munich, Germany
andreas.butz@ifi.lmu.de

ABSTRACT
This paper describes a study comparing the information recall of participants using 2D and 3D physical visualizations. Specifically, it focuses on physical bar charts and evaluates the difference between a paper-based visualization and a version built with wooden blocks. We conducted a repeated measures study involving 16 participants in which we measured the recall of information immediately after the exploration and with a delay of one week. We used questionnaires and semi-structured interviews to obtain more information about the process of recall and participants' opinions whether and how the visualizations differ in their potential for memorizing information. The results point out that participants believe to remember the 3D visualizations better, but besides the recall of extreme values the quantitative data cannot completely verify this appreciation. Furthermore the results highlight that the in the study used physical interaction techniques are not able to compensate lacking visual differentiation. One surprising finding was the strong dependency of the different data sets on the recall performance.

ACM Classification Keywords
H.1.2 User/Machine Systems: Human information proc.

Author Keywords
physical visualization; physicalization; memorability; evaluation; 2D vs. 3D.

INTRODUCTION
While digital information visualization (infovis) has a long research history, visualizations that go beyond flat digital screens have only recently started to attract attention. This development is supported by advances in digital fabrication and shape-changing displays and raises questions about how to design effective 3D visualizations for physical devices and how to study such systems to explore their potential [15].

Within this new field of research, studies have shown that 3D physical bar charts can have benefits compared to digital visualizations for information retrieval tasks [14] or the recall of information [25]. Other studies of physical visualizations investigated their suitability for visualizing personal data [19, 26], explored user interaction with physically dynamic bar charts [28] and developed tools to simplify the workflow for their design and creation [27]. However, research about physical visualizations is still at its beginning and further studies have to be conducted to investigate and compare them to alternative representation methods in order to develop strong formal evidence for their benefits [15].

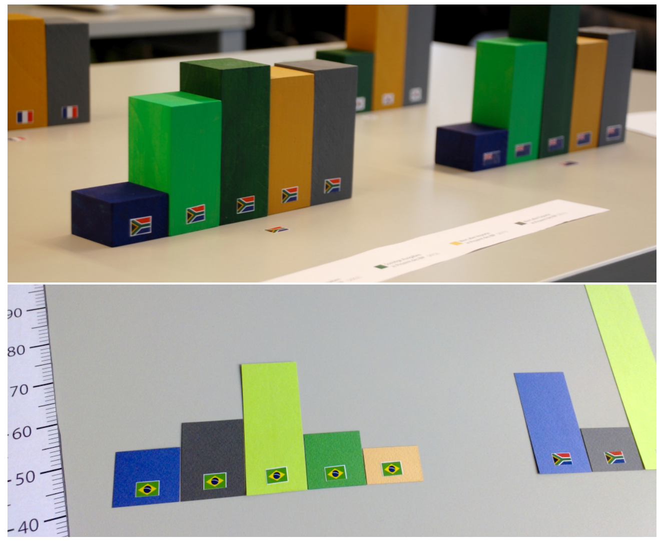

Figure 1. The physical bar charts that were used in our study. Top: 3D modality with wooden blocks. Bottom: 2D modality with paper strips.

This article attempts to provide a better understanding of which characteristics of physical bar charts influence the perception and memorization of information. In a repeated measures study with 16 participants we compared a 2D and a 3D version of token-based physical visualizations (see figure 1). Recall performance was measured by a quiz once immediately after exploration, and again after two weeks. Questionnaires and semi-structured interviews were used in order to gain further insights. The results show that characteristics that are unique to the 3D modality such as volume had a strong influence on the recall of information for extreme values. While the physical interaction with the visualizations to which we encouraged participants seemed not to have an effect on the performance, the underlying data sets had a much stronger influence than expected.

In the following sections we will first motivate our research questions and study design by discussing related work. We then report our study as well as its results and conclude with a

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. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
TEI '16, February 14 - 17, 2016, Eindhoven, Netherlands
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-3582-9/16/02$15.00
DOI: http://dx.doi.org/10.1145/2839462.2839476

discussion of implications for the design of physical visualizations and future studies.

MOTIVATION & BACKGROUND

Our work builds on research in the areas of digital and physical data representations and information recall, specifically regarding physical objects. The comparative study described in this paper investigates which characteristics of physical visualizations influence their memorability and how this potential can be used. Our main research questions are:

- Does the dimensionality of a physical visualization and the related characteristics such as weight or volume have an influence on the perception and recall of information?
- Do physical interactions such as disassembling, reassembling, grasping and lifting influence the perception and recall of information?

Data Representation & Memorability

A fundamental goal of the visual representation of data is to amplify cognition [4] and help to provide an understanding of otherwise complex data [20]. A visualization can extend human memory [33] and therefore studies about memorability are a frequently discussed topic in information visualization.

The results of a study by Borkin et al. [3] showed that visualizations are more memorable when they include objects or pictures that are easily recognized by humans. However, this does not clarify whether facts about the underlying data and information could be perceived and remembered better as the visualizations were only presented for 1 second. Guidelines for effective visualizations advise to limit a representation of data to the minimum of necessary visual elements and to avoid so-called *chartjunk* [31], which is not required to comprehend the information, or even distracts the viewer. In contrast, Bateman et al. [1] could show that *chartjunk* can enhance memorability at least for simple data sets. It is worth mentioning that the extra visual elements in their study were quite elaborate handmade illustrations and artistically embedded into the visualizations.

A study by Tory et al. [29] that compared spatializations found that the influence of additional visual elements and metaphors was hard to predict. Visual memory was more accurate when the data was perceived via dot displays and 3D landscapes compared to 2D landscapes. They also observed that the redundant encoding of data by using color and height in landscapes improved memory in comparison to color alone. Whether the additional dimension that is implied by a 3D physical visualization is rather an unnecessary embellishment or supports the perception and memorability of information remains to be investigated.

Cockburn and McKenzie [6] investigated the effectiveness of spatial memory for physical and digital arrangements of images in 2D and 3D. The users found interfaces with more dimensions less efficient, but the physical modality performed better than the on-screen modality. The first study that compared physical visualizations to their digital counterparts for low-level information retrieval tasks was done by Jansen et al. [14]. The results showed that 2D bar charts performed best for all tasks and that physical 3D bar charts performed better than the digital version. According to the authors this study was limited in the tasks and modalities used. Stusak et al. [25] took a similar approach, but focused on memorability of 2D bar charts in the digital and physical modality. The study found that participants forgot significantly less facts about maximum and minimum values within two weeks, when these were perceived from the physical bar chart. The study design allowed no comparative statements for both modalities by the participants. The authors mention that participants did hardly explore the physical visualization haptically and they suggest that further studies should investigate tasks that more clearly require physical interaction. The research agenda article about *physicalizations* by Jansen et al. [15] in addition encourages further studies that measure benefits beyond time and error and take various modalities into account, such as paper and ink.

Tangibles & Memorability

The idea of tangible interfaces is to embody digital information in physical space and therefore make use of the well-evolved human capability to sense and manipulate the physical world [13]. There are several tangible examples of visualizations. Hancock et al. [10] studied the navigation of circular tree structures via tangibles and touch. Jetter et al. [16] presented a technique that combines information visualization and tangibles to facilitate physical search queries. However, these projects concentrate on data *manipulation* with tangibles, not on data *representation* through tangibles.

O'Malley and Fraser [23] argue that there can be real benefits for learning from tangible interfaces and that carefully designed physical interactions can simplify problem solving tasks. Zuckerman et al. [36] state that the natural way to learn engages multiple senses in a constructive process. Studies by Easton et al. [9] indicate a shared abstract representation of object shape and structure for vision and haptics. Interestingly this is only the case for implicit memory tests, while explicit tests imply that the recognition system keeps track of the modality through which an object is experienced. Kerzel [18] supports the principle of intermodal information transfer with his finding that visual short-term memory is influenced by haptic perception. Similarly the experiments by Kelly et al. [17] propose that locations that are learned with different senses are represented within a common reference frame and that haptic experiences influence visually perceived memories. In addition studies showed that physical objects were recalled better than pictures, and pictures better than words [2].

This leads to the assumption that a physical representation of information could generate a more complete or detailed spatial representation in the subject's memory. However, although studies often show positive effects in the use of physical materials in education, it is not clear whether this is due to their physicality or rather to the fact that learning with physical objects typically makes use of active learning and represents information in a more salient way [30]. In addition several studies in the area of tangible interaction could not show a clear difference between the physical and virtual modality (e.g. [30, 22, 35]). However, we believe that results in the area of information visualization can differ as the previously stated studies have investigated the understanding of concepts, supposedly stored in the nondeclarative memory, but information visualization rather conveys factual information and correlations stored in the declarative memory. Our study furthermore focuses on the difference between the 2D and 3D modality and investigates whether the unique characteristics of physical objects that can be perceived haptically [21] can enhance memorability.

Memorability & Interference

While it is not possible to cope the entire research area of cognitive psychology and all theories regarding learning, memorizing and gaining knowledge, we would like to highlight some key aspects related to our study design and findings. Jean Piaget's theory of constructivist learning assumes that the assimilation of external information is dominated by internal constructions and that learning is contextual, which means that humans learn in relationship to previous knowledge (e.g. [11]).

According to the interference theory interaction between new and past learned knowledge can have a negative influence on the speed of learning and memory performance. The two main kinds of interference are *proactive* and *retroactive interference*. The former explains the forgetting of new information because of memories and knowledge that was learned beforehand (e.g. [32]). The latter describes the phenomenon when newly learned information is disrupting or hindering the recall of previous learned knowledge (e.g. [34]).

PHYSICAL VISUALIZATIONS

Based on the related studies above we decided to focus on the following aspects with our visualizations and the study:

Tap Potential & Encourage Interaction. Physical visualizations have physical characteristics such as weight, texture and hardness, which clearly distinguish them from digital visualizations. To tap this potential participants should be encouraged to explore and sense these properties. Our visualizations should support and encourage haptic interactions that go beyond touching single bars.

Enable Comparability. Finding the right alternative to a physical visualization for conducting a comparative study is difficult [15]. The data encoding, interaction techniques and physical size should be kept as similar as possible in order to achieve valid and informative results [14]. Besides digital visualizations, other baselines such as paper or whiteboards should be given consideration.

Spatial Frame of Reference. Haptic interaction can allow great flexibility in creating a visualization. This creates the danger of loosing a fixed spatial frame of reference that implies the important visual variable of spatial position. The mapping and memorization of single bars to specific categories or countries could, for example, be more difficult if the spatial arrangement varies too often. The shaping of a spatial frame of reference should be supported without limiting the possible haptic interactions.

Data Sets

We used country indicator data from the Human Development Report (HDR) as the underlying data set. The topics were *population* and *economic* trends. For both topics two data sets were extracted, each consisting of the values for six countries and five subtopics. In total we had four data sets with 30 data points each. The topics of the data sets and their complexity are in line with previous studies (e.g. [14, 25]).

Visualization Type

To minimize the influence of difficulties in reading or interpretation on the results we chose a well-known vertical bar chart as the basic visualization type. In previous studies [14, 25] physical interaction was limited to touching single bars. With this study we wanted to include further unique operations on physical visualizations such as grasping and lifting single data points. In order to achieve this we followed a token-based approach [12], where each data point is represented by an independent physical token.

Visualization Modalities

The final designs of the two visualization modalities can be seen in figure 1. We decided to concentrate on the dimensionality of physical visualizations and exclude their digital counterparts. Specifically we compared paper-based 2D bar charts with a 3D version built from wooden blocks. The single bars of both visualizations had the same size, similar colors for each category and small printed flags on them. To encourage physical interaction with the visualization we decided against labeling the bars with numeric values, but instead provided a stand-alone scale (see figure 2c&d) for reading exact values.

STUDY DESIGN

The goal of our study was to learn more about the role of the various physical characteristics of physical visualizations regarding the perception and recall of the underlying information. We also wanted to gain some understanding whether and how physical interactions influence this process. To achieve this, we first ran a within-subjects repeated measures study with the independent variables *data set* (2 levels) and *modality* (2 levels). The resulting four conditions were counterbalanced using a latin square. As the dependent variable, we measured recall performance (percentage of correct answers) once immediately after the exploration, and once with a delay of one week. In addition to this controlled experiment, we also conducted semi-structured interviews to find out about the subjects' own judgements.

Participants

We announced our study by email and social network and recruited 16 participants (5 female) with a mean age of 22.8 years. They were predominantly students of human-computer interaction (14 out of 16). Participants received a 10 Euro gift coupon for an online store.

Procedure

The study took place in an isolated and quiet room. We used a repeated measures design for the study which means that each participant worked with both visualization modalities. In

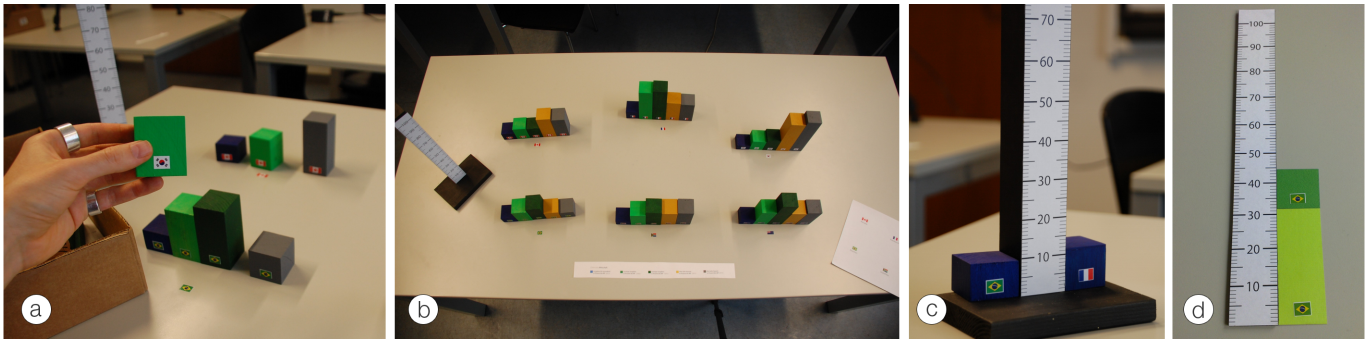

Figure 2. Participants had to (a) assemble the wooden blocks (or paper stripes) to create the (b) final visualization. Stand-alone scales for both modalities (c & d) were used in the study tasks to read and compare exact values.

order to measure the delayed recall we hold a second sessions after a gap of one week.

After a consent form and a demographic questionnaire the first session started with the *exploration phase*: one of the visualizations representing one of the data sets was presented to each participant in a counterbalanced order. Participants fulfilled several tasks with each visualization and were also asked to rate aspects such as memorability, fun and ease of use respectively on 5-point Likert scales. Participants were not told to memorize the facts of the visualization in order to test *implicit memorization*.

The *recall phase* started directly after exploring both visualizations: participants had to complete online quizzes about each data set in the same order as the presented visualizations. We advised participants to leave unknown questions blank instead of guessing the right answer. No feedback on the performance or the correct answers to the quizzes were given nor did participants know the procedure of the second session.

The second session started with a short semi-structured interview, in which participants were asked which visualization seemed to be “more present” after one week and whether they remembered any details to confirm their assumption. They then had to fill out two online quizzes again. In a second semi-structured interview participants were finally asked to what extent they had used the memorized visualizations to answer the questions and whether they had observed any differences depending on the modality.

Tasks
The tasks of the study can be split into two main categories:

Assembling the visualization. Participants sat in front of an empty table with geographically arranged flags and a legend with the order of the categories (see figure 2b). The experimenter handed over a box containing all paper strips or wooden blocks and asked the participants to assemble the visualizations. Besides the order of the categories and the rough position given by the flags on the table, participants had no instructions how to fulfill this task exactly.

Retrieving and comparing single values. The experimenter asked the participants to name the countries with the highest and lowest values for each category. Furthermore they had to give exact values for specific data points, compare specific bars (see figure 2c) and state the outcomes of summarized values (see figure 2d). Participants were encouraged to use the stand-alone scales to fulfill these tasks.

Data Collection & Analysis
In order to evaluate the recall performance and to further analyze and interpret the study results we gathered the following types of data:

Observation. The study leader observed the participants during the study and took notes, e.g., of unexpected behavior.

Video. We videotaped the entire process with two cameras from different viewing angles: a view from the front and one from the back over the participants’ shoulder. The video was mainly used to recapitulate and analyze unexpected interactions and behavior.

Questionnaires. We used pen-and-paper questionnaires to gather demographic information and the participants’ opinions on the visualizations.

Interviews. We conducted semi-structured interviews in both sessions. The goal of these interviews was to obtain more information about their process of recall and their opinions whether and how the visualizations differed in their potential for memorizing information.

Quiz. To test the recall of information participants had to fill out an online quiz directly after the exploration and with a delay of one week. It contained different question categories:

- *extreme values:* Questions about minimum and maximum values, such as “Which country had the highest GDP growth rate?”. Answers were chosen from a drop-down list and consisted of either a specific country or a specific numeric value.
- *general facts:* General questions about the underlying data, such as “In which countries did the GDP increase between 2000 and 2010?”. Answers were chosen by selecting the corresponding checkboxes.
- *summations:* Questions that included the addition of values, such as “In which countries is the addition of secondary and tertiary education higher than 60%?” Answers were chosen by selecting the corresponding checkboxes.

The quiz after one week contained several additional questions that were not part of the first one:

- *image recognition:* Pictures of the assembled physical bars for one country were displayed, but without the attached flags. Participants had to recognize the country and select the corresponding checkbox.
- *additional general questions:* General questions about the underlying data which were not asked in the first quiz.

With these questions, we wanted to investigate whether participants only remembered the answers that they had given in the first quiz or actually remembered the information from the visualizations.

RESULTS

We base all recall performance analyses and discussions on effect sizes with 95% confidence intervals as used and recommended in previous work [5]. We decided to do this because of growing concerns in several research fields regarding the limits of null hypothesis significance testing for reporting and interpreting experiment results [7, 8].

Self-Reports

After each exploration of a visualization participants filled out 5-point Likert scale questionnaires (ranging from 1=strongly disagree; to 5=strongly agree). Participants rated both visualizations as easy to use (*2D*: *MODE*=5, *M*=4.31, *SE*=0.22 / *3D*: *MODE*=5, *M*=4.69, *SE*=0.15) and inviting to interact with (*2D*: *MODE*=4, *M*=4.19, *SE*=0.16, *3D*: *MODE*=5, *M*=4.69, *SE*=0.15). This suggests that the results in the recall phase are not influenced by reading or interaction difficulties. Participants found the 3D visualization more fun to interact with (*MODE*=5, *M*=4.63, *SE*=0.13) compared to 2D (*MODE*=4, *M*=3.81, *SE*=0.31) and considered the 3D modality more memorable (*MODE*=4, *M*=4.19, *SE*=0.19) than the 2D version (*MODE*=3, *M*=3.06, *SE*=0.23).

After both sessions participants had to rank which visualization was perceived as being more "present". In the first session 11 participants ranked the 3D modality better (*2D*=3, *no difference*=2), in the second 14 participants had this opinion (*2D*=0, *no difference*=2). While almost all participants ranked the 3D modality more "present" than 2D, they had difficulties to support this assumption, e.g. by giving details about the visualization.

Recall

The effect of the modality on the overall recall results for both data sets combined revealed only a minor trend in favor of 3D. The same was true for the results of the additional questions that were only asked in the second session. We therefore ran a more detailed analysis by looking separately at the results for each data set.

Figure 3 shows the percentage of correct answers for immediate and delayed recall for both data sets. The percentage of correct answers is higher for the 3D modality, both for immediate and delayed recall, but only for the population data set. Participants seemed to have difficulties memorizing the economic data set, as the percentage of correct answers is generally much lower, independent of modality.

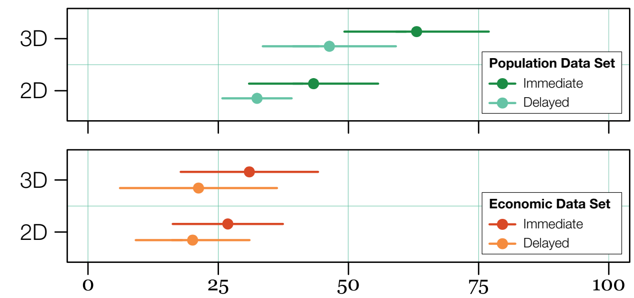

Figure 3. Percentage of correct answers for *immediate* and *delayed* recall for the *population* and *economic* data set (with 95% CIs).

Figure 4 looks only at the population data set and shows the percentage of correct answers about extreme values. The split into minima (lowest bars) and maxima (highest bars) reveals that maximum values in general could be remembered better in the 3D modality. While Figure 4 shows only the results for the population data set, the economic data set shows a similar trend but much less distinct.

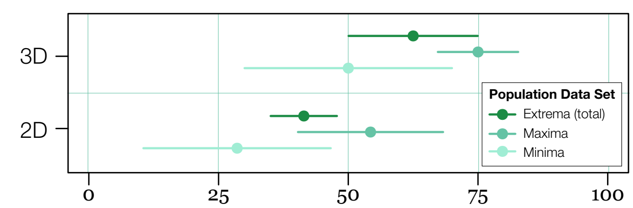

Figure 4. Percentage of correct answers for *immediate* recall of the *population* data set, divided into maxima, minima and total extrema (with 95% CIs).

Figure 5 shows a different view on the performance for the economic data set. The results are grouped by participants who had the economic data set encoded in the first or second visualization for both modalities. The results show that the order in which the visualizations for both data sets was handed out had a clear influence on the recall of information from the economic data set. Surprisingly, this effect could not be observed for the population data set.

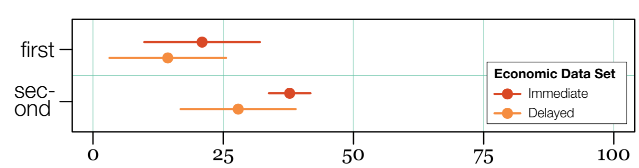

Figure 5. Percentage of correct answers for *immediate* and *delayed* recall of the *economic* data set, depending on whether it was encoded in the first visualization that was presented or the second (with 95% CIs).

Discussion

In this section we will discuss the observed strong influence of the data set on the recall performance, and also discuss our other findings in light of our initial questions regarding the role of unique physical characteristics and physical interactions for the perception and recall of information.

The Role of the Data Set

There is a clear difference of the recall results regarding the two data sets. While the population data set indicates an advantage for the 3D modality in both quizzes, the economic data set shows no such difference. In the semi-structured interviews, we learned that most participants bore no relation to economic topics and stated that the data was too abstract and less interesting. This might have been a side effect of our choice of HCI, i.e., computer science students as the subjects and could be explained with the presumption of contextual learning and the constructivism theory. In contrast, participants stated that the population data set consisted of *"concrete and easily imaginable categories"*. These aspects seemed to have a relevant influence on the perception and memorization of the visualization and the underlying data

The quiz results further indicate that the information gained from a visualization encoding the economic data set is "overwritten" by the following information from the visualization of the population data set, but not vice versa. This is in a certain sense consistent with the phenomena described by the interference theory. Questions about the economic data set were left blank more often than questions about the population data set, which suggests that participants did not remember the information wrong, but had difficulties in remembering it at all.

A detailed look at the range of values in both data sets provides an additional potential explanation: The distance between extreme values and their neighbors was higher in the population data set, and the resulting more distinct visual differences could potentially also increase the potential for memorization.

As a critical aspect, it should be mentioned that visualizations are often used to make complex data more accessible and it seems that the chosen visualizations were not able to fulfill this purpose well for the economic data set. In general, our study shows that the data set has a strong influence on the recall performance and therefore the choice of test data sets should be well-considered, especially for studies about memorability. We furthermore suggest, that future studies should consider the role of the underlying data more carefully, e.g., by integrating the exploration and analysis of a visualization into a main problem solving tasks, where comprehension of the data is required and meaningful.

The Role of Interaction

As stated before the visual difference between extremes and the neighboring values was higher in the population data set. This also had an effect on the interaction, as participants directly compared two bars in the economic data set on average almost six times to identify the higher or lower one. This was uncommon for the population data set, in which such comparisons were made only two times on average. However, the results imply that the higher number of interactions could not compensate the lack in visual differentiation.

The physical interactions during the study consisted of two main parts: assembling the overall visualizations from single bars and measuring or comparing single bars to retrieve values. While two participants stated that the manual assembly helped them two remember the categories and countries, most participants described it as "mechanical" and tried to finish it as fast as possible. The same applies for the measuring interaction, where "mechanical" was again an often used term. Some participants also declared that they had assumed that a fast and unproblematic interaction with the bars was the focus of the study and that the actual data was rather unimportant.

Further statements of participants in the interviews in the second session emphasize that the physical interactions seemed rather unnecessary and were performed almost automatically. Four participants for example did not remember any use of the stand-alone scale until an explicit question about it. Another participant mentioned that he used the scale very rarely, which is surprising considering that the measuring was a main part of the study tasks.

This shows that the design of effective and meaningful physical interactions for physical visualizations is challenging and that the chosen approach to "enforce" interaction was not productive. We believe that especially the high repetition of the same, rather artificially encouraged, physical interaction of assembling and moving paper stripes or wooden blocks was too common to generate actual benefits. We suggest that future studies should offer a broader range of possible physical interactions and try to emphasize particular insights into the data set with the rare use of specific interactions.

The Role of Modality

The results show that participants could remember information better when they perceived it with the 3D modality. This is especially the case for extreme values, which confirms the results of the study by Stusak et al. [25]. They argue that the results can be attributed to the *"vivid physical height of the bars"*, which is in line with the findings of Scott [24] that object names are less memorable than pictures, because they are less distinctive. Our study furthermore reveals that the distinctiveness is based on the 3D modality and not mandatory for physical visualizations, as the 2D modality in our study led to similar results as the digital bar charts in the study by Stusak et al. [25].

We believe that the additional characteristics of the 3D wooden blocks such as the top and side faces or the volume contribute to the distinctiveness and explain the benefits compared to the 2D paper stripes. One participant furthermore stated that the 3D modality replaced the 2D modality completely in her impression. None of the participants said that they did notably perceive the weight of a single wooden block or could make use of the sensed information. This is probably the case because humans have evolved an expectation and "sensation" about the weight of common physical objects.

As we do not believe that one of our two visualization modalities is more novel or ordinary than the other, we assume that we can rule out a novelty effect, which is often claimed in the case of physical visualizations.

SUMMARY AND FUTURE WORK

With the goal to investigate the influence of the particular properties of physical visualizations, such as weight, volume or texture, on their memorability, we have run a comparative study between 2D paper bar charts and 3D bar charts built from wooden blocks. We used two different data sets and evaluated recall performance immediately after exploration and with a delay of a week. The study was accompanied by semi-structured interviews and questionnaires regarding the interactions done in the main task and the subjective assessment of memorability.

When looking at the measured recall performance, we found that with one data set, the 3D visualization caused a much better recall in general than the 2D variant. With the other data set, however, this effect was not present. A closer look at the interview results revealed, that a main difference between these data sets was, that participants could relate well to one of them, while they found the other one too abstract and not interesting. While the overall recall performance for the uninteresting data set was lower (as can be expected), also the difference between the 3D and 2D visualizations in terms of recall disappeared. This suggests that the particular properties of spatiality and tangibility can show stronger benefits when the underlying data set is also comprehensible and interesting. In other words: The better the mind can grasp the information behind a physical visualization, the more will grasping with the hands help.

As stated in the discussion, we could not find an influence of the physical factor weight, possibly because the weight of objects in our studies was totally conform with typical expectations. Further studies could try to investigate whether breaking these expectations could generate additional benefits. However, we believe that this effect might be limited regarding scalability and could, for example, be used to highlight a single data point. A different approach could be to use materials with a higher density to increase the absolute difference in weight. In the area of actuated physical visualizations, the combination with weight changing user interfaces could, for example, enable the adaption of the weight of single data points and, therefore, modify the required physical exertion to lift a particular data point.

REFERENCES

1. Scott Bateman, Regan L. Mandryk, Carl Gutwin, Aaron Genest, David McDine, and Christopher Brooks. 2010. Useful Junk?: The Effects of Visual Embellishment on Comprehension and Memorability of Charts. In *CHI '10*. ACM, New York, NY, USA, 2573–2582. DOI: http://dx.doi.org/10.1145/1753326.1753716
2. William Bevan and Joseph A Steger. 1971. Free recall and abstractness of stimuli. *Science* 172, 3983 (1971), 597–599. DOI: http://dx.doi.org/10.1126/science.172.3983.597
3. Michelle Borkin, Azalea Vo, Zoya Bylinskii, Phillip Isola, Shashank Sunkavalli, Alfonso Oliva, Hanspeter Pfister, and others. 2013. What makes a visualization memorable? *Visualization and Computer Graphics, IEEE Transactions on* 19, 12 (2013), 2306–2315. DOI: http://dx.doi.org/10.1109/TVCG.2013.234
4. Stuart K Card, Jock D Mackinlay, and Ben Shneiderman. 1999. *Readings in information visualization: using vision to think*. Morgan Kaufmann.
5. Fanny Chevalier, Pierre Dragicevic, and Steven Franconeri. 2014. The not-so-staggering effect of staggered animated transitions on visual tracking. *Visualization and Computer Graphics, IEEE Transactions on* 20, 12 (2014), 2241–2250. DOI: http://dx.doi.org/10.1109/TVCG.2014.2346424
6. Andy Cockburn and Bruce McKenzie. 2002. Evaluating the Effectiveness of Spatial Memory in 2D and 3D Physical and Virtual Environments. In *CHI '02*. ACM, New York, NY, USA, 203–210. DOI: http://dx.doi.org/10.1145/503376.503413
7. Geoff Cumming. 2013. The new statistics why and how. *Psychological science* (2013), 0956797613504966.
8. Pierre Dragicevic. 2015. *HCI Statistics without p-values*. Research Report RR-8738. Inria. 32 pages. https://hal.inria.fr/hal-01162238
9. Randolph D Easton, Anthony J Greene, and Kavitha Srinivas. 1997. Transfer between vision and haptics: Memory for 2-D patterns and 3-D objects. *Psychonomic Bulletin & Review* 4, 3 (1997), 403–410. DOI: http://dx.doi.org/10.3758/BF03210801
10. Mark Hancock, Otmar Hilliges, Christopher Collins, Dominikus Baur, and Sheelagh Carpendale. 2009. Exploring Tangible and Direct Touch Interfaces for Manipulating 2D and 3D Information on a Digital Table. In *ITS '09*. ACM, New York, NY, USA, 77–84. DOI: http://dx.doi.org/10.1145/1731903.1731921
11. George Hein. 1991. Constructivist learning theory. *Institute for Inquiry.* (1991). http://www.exploratorium.edu/ifi/resources/constructivistlearning.html
12. Samuel Huron, Yvonne Jansen, and Sheelagh Carpendale. 2014. Constructing visual representations: Investigating the use of tangible tokens. *Visualization and Computer Graphics, IEEE Transactions on* 20, 12 (2014), 2102–2111. DOI: http://dx.doi.org/10.1109/TVCG.2014.2346292
13. Hiroshi Ishii and Brygg Ullmer. 1997. Tangible Bits: Towards Seamless Interfaces Between People, Bits and Atoms. In *CHI '97*. ACM, New York, NY, USA, 234–241. DOI:http://dx.doi.org/10.1145/258549.258715
14. Yvonne Jansen, Pierre Dragicevic, and Jean-Daniel Fekete. 2013. Evaluating the Efficiency of Physical Visualizations. In *CHI '13*. ACM, New York, NY, USA, 2593–2602. DOI: http://dx.doi.org/10.1145/2470654.2481359
15. Yvonne Jansen, Pierre Dragicevic, Petra Isenberg, Jason Alexander, Abhijit Karnik, Johan Kildal, Sriram Subramanian, and Kasper Hornbæk. 2015. Opportunities and Challenges for Data Physicalization. In *CHI '15*. ACM, New York, NY, USA, 3227–3236. DOI: http://dx.doi.org/10.1145/2702123.2702180

16. Hans-Christian Jetter, Jens Gerken, Michael Zöllner, Harald Reiterer, and Natasa Milic-Frayling. 2011. Materializing the Query with Facet-streams: A Hybrid Surface for Collaborative Search on Tabletops. In *CHI '11*. ACM, New York, NY, USA, 3013–3022. DOI: http://dx.doi.org/10.1145/1978942.1979390

17. Jonathan W Kelly, Marios N Avraamides, and Nicholas A Giudice. 2011. Haptic experiences influence visually acquired memories: Reference frames during multimodal spatial learning. *Psychonomic bulletin & review* 18, 6 (2011), 1119–1125. DOI: http://dx.doi.org/10.3758/s13423-011-0162-1

18. Dirk Kerzel. 2001. Visual short-term memory is influenced by haptic perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27, 4 (2001), 1101–1109. DOI: http://dx.doi.org/10.1037/0278-7393.27.4.1101

19. Rohit Ashok Khot, Larissa Hjorth, and Florian 'Floyd' Mueller. 2014. Understanding Physical Activity Through 3D Printed Material Artifacts. In *CHI '14*. ACM, New York, NY, USA, 3835–3844. DOI: http://dx.doi.org/10.1145/2556288.2557144

20. Claire Knight. 1998. *Visualisation for program comprehension: information and issues*. Citeseer.

21. Susan J Lederman and Roberta L Klatzky. 2009. Haptic perception: A tutorial. *Attention, Perception, & Psychophysics* 71, 7 (2009), 1439–1459. DOI: http://dx.doi.org/10.3758/APP.71.7.1439

22. Paul Marshall, Peter C.-H. Cheng, and Rosemary Luckin. 2010. Tangibles in the Balance: A Discovery Learning Task with Physical or Graphical Materials. In *TEI '10*. ACM, New York, NY, USA, 153–160. DOI: http://dx.doi.org/10.1145/1709886.1709914

23. Claire O'Malley and Danae Stanton Fraser. 2004. Literature Review in Learning with Tangible Technologies. Research report. (2004). https://telearn.archives-ouvertes.fr/hal-00190328

24. Keith G Scott. 1967. Clustering with perceptual and symbolic stimuli in free recall. *Journal of Verbal Learning and Verbal Behavior* 6, 6 (1967), 864–866.

25. Simon Stusak, Jeannette Schwarz, and Andreas Butz. 2015. Evaluating the Memorability of Physical Visualizations. In *CHI '15*. ACM, New York, NY, USA, 3247–3250. DOI: http://dx.doi.org/10.1145/2702123.2702248

26. Simon Stusak, Aurélien Tabard, Franziska Sauka, Rohit Ashok Khot, and Andreas Butz. 2014. Activity sculptures: Exploring the impact of physical visualizations on running activity. *Visualization and Computer Graphics, IEEE Transactions on* 20, 12 (2014), 2201–2210. DOI: http://dx.doi.org/10.1109/TVCG.2014.2352953

27. Saiganesh Swaminathan, Conglei Shi, Yvonne Jansen, Pierre Dragicevic, Lora A. Oehlberg, and Jean-Daniel Fekete. 2014. Supporting the Design and Fabrication of Physical Visualizations. In *CHI '14*. ACM, New York, NY, USA, 3845–3854. DOI: http://dx.doi.org/10.1145/2556288.2557310

28. Faisal Taher, John Hardy, Abhijit Karnik, Christian Weichel, Yvonne Jansen, Kasper Hornbæk, and Jason Alexander. 2015. Exploring Interactions with Physically Dynamic Bar Charts. In *CHI '15*. ACM, New York, NY, USA, 3237–3246. DOI: http://dx.doi.org/10.1145/2702123.2702604

29. Melanie Tory, Colin Swindells, and Rebecca Dreezer. 2009. Comparing dot and landscape spatializations for visual memory differences. *Visualization and Computer Graphics, IEEE Transactions on* 15, 6 (2009), 1033–1040. DOI:http://dx.doi.org/10.1109/TVCG.2009.127

30. Lara M Triona and David Klahr. 2003. Point and click or grab and heft: Comparing the influence of physical and virtual instructional materials on elementary school students' ability to design experiments. *Cognition and Instruction* 21, 2 (2003), 149–173. DOI: http://dx.doi.org/10.1207/S1532690XCI2102_02

31. Edward R Tufte and PR Graves-Morris. 1983. *The visual display of quantitative information*. Vol. 2. Graphics press Cheshire, CT.

32. Benton J Underwood. 1957. Interference and forgetting. *Psychological review* 64, 1 (1957), 49. DOI: http://dx.doi.org/10.1037/h0044616

33. Colin Ware. 2012. *Information visualization: perception for design*. Elsevier.

34. Erica L Wohldmann, Alice F Healy, and Lyle E Bourne Jr. 2008. A mental practice superiority effect: less retroactive interference and more transfer than physical practice. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34, 4 (2008), 823. DOI: http://dx.doi.org/10.1037/0278-7393.34.4.823

35. Zacharias C Zacharia and Georgios Olympiou. 2011. Physical versus virtual manipulative experimentation in physics learning. *Learning and Instruction* 21, 3 (2011), 317–331. DOI: http://dx.doi.org/10.1016/j.learninstruc.2010.03.001

36. Oren Zuckerman, Saeed Arida, and Mitchel Resnick. 2005. Extending Tangible Interfaces for Education: Digital Montessori-inspired Manipulatives. In *CHI '05*. ACM, New York, NY, USA, 859–868. DOI: http://dx.doi.org/10.1145/1054972.1055093