Scaling Data Physicalization – How Does Size Influence Experience?

Irene López García
irene.lopez.garcia@uni-weimar.de
Bauhaus-Universität Weimar
D-99423 Weimar, Germany

Eva Hornecker
eva.hornecker@uni-weimar.de
Bauhaus-Universität Weimar
D-99423 Weimar, Germany

ABSTRACT

Given the material nature of data physicalization, their creators need to make many design decisions, including material choices and scale. Our study explores the impact of scale in physicalization, motivated by the assumption that size can affect user experience. We created two different physicalizations (for the same dataset) in three sizes each, and evaluated the resulting six objects with a questionnaire approach and interviews. Our findings highlight that scale needs to be chosen wisely given its impact on representation legibility (ease of viewing and understanding) and affordances for interaction. We discuss factors to take into account when designing large-scale physicalizations and in further research on the potential role of scale in physicalizaton. In particular, we argue that for large-scale physicalizations, scale should matter and communicate meaning, for instance, supporting an intuitive understanding of magnitudes, or a specific experience. Thus, scale needs to be an explicit design decision, that interacts with other design parameters.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; **Empirical studies in HCI**; **Empirical studies in visualization**;

KEYWORDS

InfoViz; user experience; user study; Design; Scale; Materiality; legibility; affordances for action

1 INTRODUCTION

If we think of data representation, we normally think of visualizations. But what if data moved off the screen, with 3D physical shape and materiality, into the world, so we could touch it, hold it in our hands, feel it – or navigate the data by moving around it? This is the field of data physicalization, which has become a shared

topic of interest across InfoViz, HCI, but also in Design and the Arts [11, 12, 14].
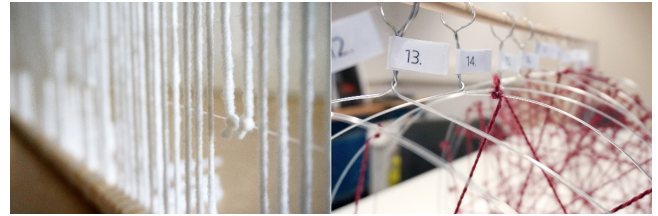

Figure 1: Close-up on two of the models made for the study.

Given the material nature of physicalizations, which do not have the inherent fluidity of purely digital representations, their creators need to make many conscious design decisions, including material choice and scale/size (cf. [20]). Parameters such as materiality, shape, color, and so on may all influence appearance and thus legibility of data physicalizations, as well as user experience. Indeed, some physicalizations are intentionally large, sometimes approaching architectural scale. In the arts, overall size of an object/installation is used as an expressive instrument.

As physical representations share the same physical space with users, unlike objects on a screen, it is to be assumed that these will be experienced differently depending on how our body relates to them. This is inspired by phenomenological discourse in architecture on the role of the body and kinesthetics in our experiences of space and physical structures, and neuropsychological evidence on enactive perception in architectural experience [26, 35, 49]. We know from the Arts that scale influences the viewing experience [4, 37]. The role of scale on experience is also an explicit concern in media architecture [1, 33]. This motivates the question whether a larger representation is experienced differently, impacting on an audience's reaction to its message. Moreover, smaller data physicalizations may be hand-held and can be brought to the eye and turned around, while large-scale constructions may require the viewer to navigate around the object. Thus, size impacts affordances for action, e.g. the ability to perceive objects (viewing angle, ability to navigate) or to manually handle them. This may further influence the user experience and the ability to interpret a representation. We here begin to study these issues, with size (or scale) as a design factor for a comparative study.

We developed two alternative data representations based on a data set and realized both in three sizes, from hand-held to 1m height. We aimed for a fairly neutral design that would not be too artistic, but still unlike traditional visualizations. The six resulting constructions were located at public places and 85 members of

the public were recruited to explore one object and respond to a questionnaire. For deeper insight, another eight people experienced one design in all three sizes in a semi-structured interview.

We contribute to a deeper understanding of physicalization design and the factors influencing user experience. Our experiences highlight that it is difficult to only alter size as an independent variable, as scaling a material design often requires changes in materials and construction. This is an important lesson for future work. We found that smaller representations tend to be easier to 'take in' visually and to manually interact with, and this inclined participants to prefer them over larger versions. Larger sizes were rated as impressive and better suited for public settings. We discuss that viewability needs to be considered in design of physicalization, and that large scale should be purposefully chosen to create extra value from scale. We further provide suggestions for future research on the role of scale and experience.

2 BACKGROUND

We introduce data physicalization and its characteristics, benefits and limitations identified so far. We then focus on the role of scale, discussing prior work from adjacent areas.

2.1 Data Physicalization

Jansen et al. [24] define 'physicalization' (analogous to 'visualization') as 'physical artefacts whose geometry or material properties encode data'. This is an emerging topic [12, 13, 54] across Information Visualization, HCI, ambient displays, tangible interaction, the arts and design [3, 9, 15, 18, 21, 23, 24, 27, 31, 32, 41]. Current research [13, 24] begins to systematically investigate how physicalizations are understood by people, and to establish design principles.

Similar to tangible interfaces [42], physicalizations do not have the infinite flexibility of digital representations with different data views. But their physicality lends physical presence and persistence, supporting motor memory, and affordances for interaction [42]. Benefits are seen in greater active and multimodal sensory perception than with traditional 2D-visualizations [21, 24, 46]. Physicalizations leverage natural 3D stereo-vision with perfect depth of field [53], and enable direct touch. This supports cognition, in particular spatial cognition (via active exploration) when tangible objects are touched, held, handed around, watched at eye level. They thus can support data understanding for laypeople, conveying statistics in a less abstract way. The ability to touch 'data' can improve memory of information, in particular for extreme items [44], increase engagement [9, 18, 27], and efficiency in retrieving information [23]. Moreover, physicalizations are ideal for ambient data displays and can be embedded in the context they provide information on [56]. Overall, current work highlights advantages for interactive exploration, but also identifies limitations in terms of the ease, for instance, of reading accurate values and ambiguities as to what aspect of a shape is relevant [25].

Because of these benefits, physicalization is often used in data communication and education, especially to raise public awareness of topics [13], either as demonstration medium, or as standalone representation. Many physicalizations in museums, galleries and public places rely on the audience to see, examine, and interpret them. Examples include the walkable landscape of 'Blood Swept Land and Sea of Red' [11], Adrien Segal's ephemeral, melting ice sculpture 'Grewingk Glacier' [41], Orbacles [30], etc. Many other examples can be found on the data physicalization wiki [54].

Most current work in physicalization translates conventional representations, such as bar charts, into 3D materiality, and evaluates effectiveness and efficiency of conveying information [23, 25, 44]. But vande Moere (who proposed the notion of 'data sculpture' [31, 59]) suggests there may be unique qualities for how people engage with data, affording "communication of much richer, multi-faceted and interpretative meanings than just quantitative values" [32]. UX aspects of physicalization remain under-studied. Wang et al. [53] also argue that speed of analysis and accuracy should not be the only parameters for assessing data representations, given that many physicalizations from the arts and design tradition intend to raise curiosity, provide surprise, or create an emotional connection and impact. This motivates our work here, which investigates whether physicalization size influences user experience. In particular, it would be interesting whether larger size can convey a stronger 'message', with more emotional or cognitive impact.

2.1.1 Small (Hand-Held) versus Large Physicalizations. With physical objects, size is perceived by the viewer in relation to their own body [50, 51]. Many data physicalizations developed within HCI and in design tend to be hand-held size, and often are 3D-fabricated (see figure 2, middle and right). Work includes data sculpture necklaces depicting air quality, 3D-printed chocolate of physical activity logs, Lego for activity logging, 3D-printed bar charts of country indicators or of physical activity data [45]; and the emerging practice of data knitting includes scarfs and larger creations such as blankets and throws (see http://dataphys.org/list). Such physicalization may also be interactive and involve other modalities [7, 14, 52, 52]. Larger constructions for data physicalization are rarer and tend to be actuated 3D-surfaces, such as MIT's inFORM project and its successors [15, 34], which can be described as furniture-size.

In the Arts, large physicalizations are far more common, including objects at architectural scale, room-sized setups, objects of super-human size, and furniture-sized objects. The installation 'Blood Swept Lands and Seas of Red' [11] at the Tower of London represented fallen soldiers in WW1 with a sea of ceramic poppies. 'Orbacles' [30] are 30-ft. steel sculptures showing the impact of climate change on bird species. 'Mount Fear' [39] shows an elevation map of London crime rates as a walkable room-sized sculpture (figure 2 left). 'Of all the People in all the World' [8] maps each person with a grain of rice, assembled to heaps, to convey statistics of huge numbers.

2.2 The Role of Scale for Data Physicalization

Scale refers to the overall size of the entire object. With physicalizations, scale is an explicit, immutable design decision (unless extreme forms of shape-change are used). Despite it being a central design decision, the role of scale has rarely been investigated in data physicalization research so far. We therefore also look at related fields, which study and discuss issues related to scale or size.

2.2.1 Scale in Physicalization and Shape-Change. In the emerging field of shape-changing interfaces [2, 10, 15, 34, 38] (which overlaps with data physicalization), shape-change is often used to create a
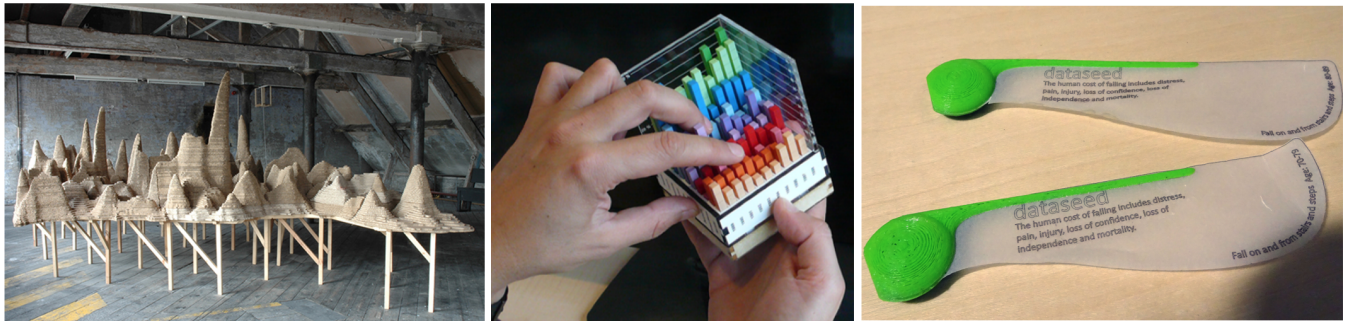
Figure 2: Examples of data physicalization. From room-sized: Mount Fear shows London crime rates as a large elevation map (© Abigail Reynolds, with permission) [39], to small: country indicators over time (from: [24]). These can also be interactive: The Dataseeds by Dulake and Gwilt [14] translate data on falls at age 50-89, where surface area of the wing is determined by the data, which influences spin and fall speed when dropped (own photo).

3D moving data representation. So far, it has treated change of scale or size mostly as a technical problem, necessitated by the need to depict changes of variables, often relying on pin-based displays. In their overview of grand challenges in shape-change research, Alexander et al. [1] mention investigating the "impact of scale and resolution" on how well users can detect and distinguish changes and on user experience as open research challenges.

Stusak [45] mentions scalability as a challenge when the value range for data to be 3D-printed is not known beforehand, or if there are extreme values (variation in small values as well as larger variations). In such cases, the scale of a 3D-print has implications for which part of the data can be more easily inspected (given physicalizations cannot be zoomed in and out). Jansen et al. [25] conducted what is apparently so far the only systematic study of the impact of scale/size, investigating effects on the ability to read and compare data values. Their findings highlight that change in 3D size is not always perceived linearly. Participants were successfully able to compare the height of 3D bars (and comparing volume of 3D squares is doable). But with spheres of different sizes, volume and diameter change at different levels. The first problem this creates is that it is unclear which of the two to 'read' as data parameter. Moreover, while participants were able to compare diameter (although not as good as for the bars), error rates in comparing volume were high. Thus, when encoding data in 3D objects, care needs to be taken in the mapping of object parameters to be clear and easy to decipher and compare. Both publications focus on relative size within one physicalization and the influence on legibility, rather than on wider experience effects of its overall scale.

2.2.2 Scale as an Explicit Design Element in the Arts and Architecture. Disciplines outside HCI have a longer history of considering scale, in particular in relation to the human body [4]. Both the Arts and Architecture use scale as an explicit design element.

In the Arts, scale is known to influence viewing experience, and artists often choose extreme scale as an 'expressive element', shrinking the viewer to change their relationship to the depicted object, or miniaturizing the latter to invite intimate inspection [37], contrasting or exaggerating scale as a design statement [4]. Scale as a relational element has the viewer's body as reference [4] (see [50, 51]) and, according to Puric, "directs attention to the relation between a work and location or place" [37]. The latter is related to what could be termed situatedness [42] or embeddedness [56].

In architecture, phenomenological discourse on the role of the body and of kinesthetics in our experience of space and physical structures emphasizes the relation between body and architecture (and thus implicitly of scale) and the implications for perception and action. Pallasmaa writes: "Understanding architectural scale implies the unconscious measuring of the object or the building with one's body and of projecting one's body scheme into the space in question" [35] and notes that our experience is shaped by bodily reaction [35]. Moreover, based on evidence from neuropsychology, Jelić et al. [26] argue "that we engage with architecture through embodied action" and that sculptural and architectural objects are perceived and understood in terms of the opportunities for action they provide. Thus, physical size can have perceptual and affective impact, for instance, large structures may be experienced as impressive or as towering, or as spacious and generous.

Media-architecture, a field between architecture and HCI, also takes account of scale, for instance regarding micro and macro level interaction, where the macro level tends to refer to urban places and city scale [1, 33].

2.2.3 Physical Manipulation and Navigation. Scale is a relational element with the user's body as reference point [4, 26, 35, 37, 50, 51]. Therefore we can distinguish between hand-held size, medium scale (too large to be hand-held, up to small furniture size), human-size, and architectural scale. Many benefits of physicalization discussed earlier (such as interaction affordances, haptic memory, active exploration) apply for a scale that can be physically manipulated and rotated, that is, for hand-held-sized objects. The dataseeds [14] (figure 2 right) only reveal spin and fall speed when dropped, and thus rely on being small enough to be picked up. With large physicalizations (at furniture, body or architectural scale), the viewers themselves need to navigate around the object. Physical scale thus changes the affordances for action and perception.

Many known limitations of 3D visualisations do not hold for physicalizations. Wang et al. [53] argue that, regardless of size, different to screen-based 2D and 3D-visualizations, physicalizations are true 3D structures. They allow for stereo-vision and can be explored from multiple perspectives through an observer's movement

(rotating the object or navigating around it), and can potentially engage all senses [53]. This includes the tactile sense. Actively touching an object reveals a multitude of properties, and memory of the experience can endure for days [28]. Beyond added information, touch creates an emotional connection – people like touching [17, 18, 53]. Data sculpture and tangible forms of data representations have been found to stimulate discussion and support collaborative information discovery [9, 18]. Khot et al. [27] observed that when viewing and handling 3D-printed objects of their own physical activity data, people gained a better sense of it and had stronger emotional connections than with visualizations.

While the impact of handling physical objects has begun to be explored, the wider influence of scale has been mostly ignored. It is open whether people also want to touch larger-scale objects in order to explore them; besides, with truly large objects, it will be impossible to touch all of it. Similarly, affordances for navigation (and its effort) change with size. We will return to these questions in our discussion, as our study found effects on the ability to view, navigate, and physically handle the physicalizations.

2.2.4 The Effect of Screen Size . Scale as a factor has only been rarely addressed in HCI research – with purely virtual, digital representations this is not a pressing issue, apart from questions about the ideal size of hand-held devices (see e.g. [29]). A cluster of related prior HCI research systematically investigates the effect of screen size for interactive visualization.

Even with identical viewing angle (same space taken up in visual image), larger screens (set further away) were found to result in better spatial orientation and higher sense of presence [47]. Furthermore, Ball et al. [6] found that access to a large display appears to make people rely more on bodily movement in navigation strategies, and that physical navigation tends to be more efficient than virtual navigation (zooming etc.), as it retains overview while viewing detail. The kinesthetic and proprioceptive cues generated by physical navigation support spatial recall and memory [3, 48, 58], in particular if items are persistent in space [48]. Evidence is not consistent, though. Jakobsen and Hornbæk [40] found that visual search speed for targets on large screens was impaired if participants could not see the entire space at once. People preferred a screen that fits their field of view. Physical, direct interaction has further been found to decrease performance, being slower than indirect action (mouse or pad), albeit supporting spatial memory; moreover, the higher effort resulted in lower user satisfaction [58].

In research on screens, larger scale is usually utilized for higher resolution, more detail, focus+context etc., with dynamic and interactive visualizations. It is unclear whether findings generalize to physicalization, in particular as they do not relate to the representation itself, but to its conveyor, the screen. Nevertheless, there is likely overlap, and some findings, e.g. physical navigation and manipulation supporting spatial memory, coincide with findings for data physicalization.

3 TWO PHYSICALIZATIONS - DESIGN AND SCALING

We created two designs for the same underlying data set. Two designs would allow for wider insight and comparison. As data to be represented, a list of documented extinct species from Wikipedia

was chosen [55]. The topic was selected as it is of public interest and might create an affectiv response. While not a comprehensive record from what is estimated to be a much larger number [36], this list includes those species well documented. The last thousand years were extracted as data, and tabulated as numbers. The timeline was then segmented into decades and centuries.

Yau proposes that good data visualization should allow users "to see trends, patterns, and outliers that tell you about yourself and what surrounds you" [57, p.xi]. A core design rationale was thus clarity and readability, with a fairly neutral and abstract design (that directly shows the data). Nevertheless, this should be aesthetically different from traditional charts. We refer to the two representations as Design A (white threads) and B (the hoops). We describe the general design and then the design rationale, before detailing how these were fabricated.

Design A (figure 3) depicts a fairly traditional bar chart timeline in 10-year steps. Each decade (1020s to 2010s) is represented by a white vertical thread, laid chronologically from left to right in a rectangular frame. Horizontal translucent nylon threads hold these in place, for more stability. The chart's scales are inscribed in the frame (or attached, see photos): decades on the X-Axis, and numbers of documented extinctions in units for the Y-Axis. The length of a thread indicates the amount of extinctions: a shorter thread (from top-down) means more species went extinct. The data is shown as gaps in the weave, thus using emptiness (missing white thread) as a metaphor for the loss of biodiversity.

Design B (figure 4) segments the data in centuries from 12th to 21st. Each century is depicted by a hoop (ten in total), hung from a supporting structure (a wooden stick). Red threads are laid randomly across each hoop, where each thread represents a species gone extinct. The idea is that each hoop provides a snapshot (what happened in a century), but one can look at them from the side in a summative view (figure 5). The increasing number of threads allows to see how extinction is not just a matter of a certain time period, but the loss in biodiversity is additive and effects continue into the future. The metaphorical value of this design lies in the red thread reminding of the lifeblood of species. Another aspect is that despite being a static representation, its segments can be moved and inspected, to see and compare particular snapshots.

3.1 Design Rationale

Even if this is not immediately visible, Design A is a fairly standard data representation (basically a flipped bar chart) and does not make use of 3D. With Design B, we aimed for a less conventional, innovative and curiosity-raising representation that would extend into 3D, but should nevertheless be abstract as well as directly encode the data so it is easy to read once the mapping is known.

Our data source covers 100 decades. Design A shows the number of extinct species as a gap in the thread for every 10-year period, which allows to see the time lapse as a 'broken' weave. The design decisions for Design B of movable pieces in the third dimension and of representing each extinct species as a thread called for less granular data aggregation. 100 hoops would be far too unwieldy as a construction (taking up a lot of space, hard to build), as well as getting too complex, thus not lending itself to exploration and seeing evolution through time. Thus, we decided to aggregate the
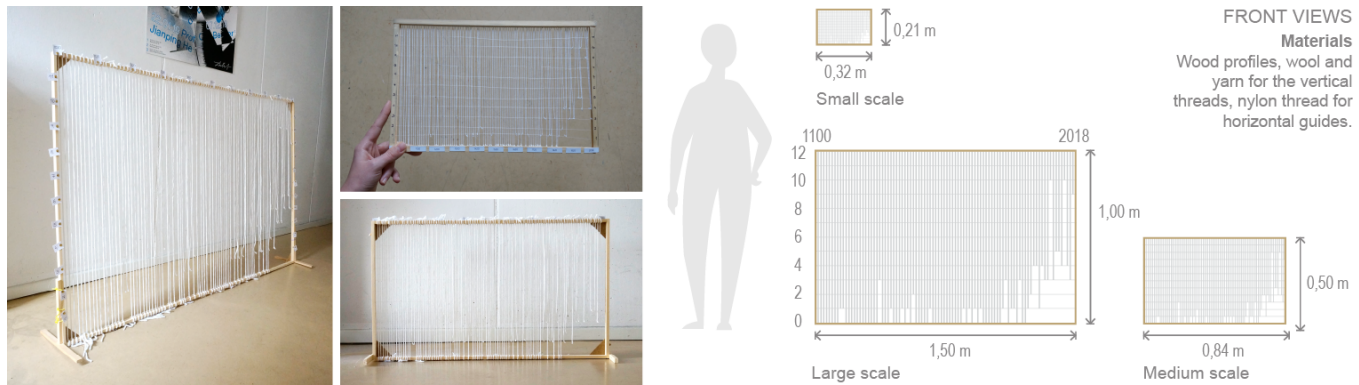
Figure 3: Design A. Photos of the largest (1,5x1m) version, the medium (0,84x0,5m) and smallest version (0,32x0,21m) (being held up). For better comparison, the diagram on the right shows the designs in relation to a human body.
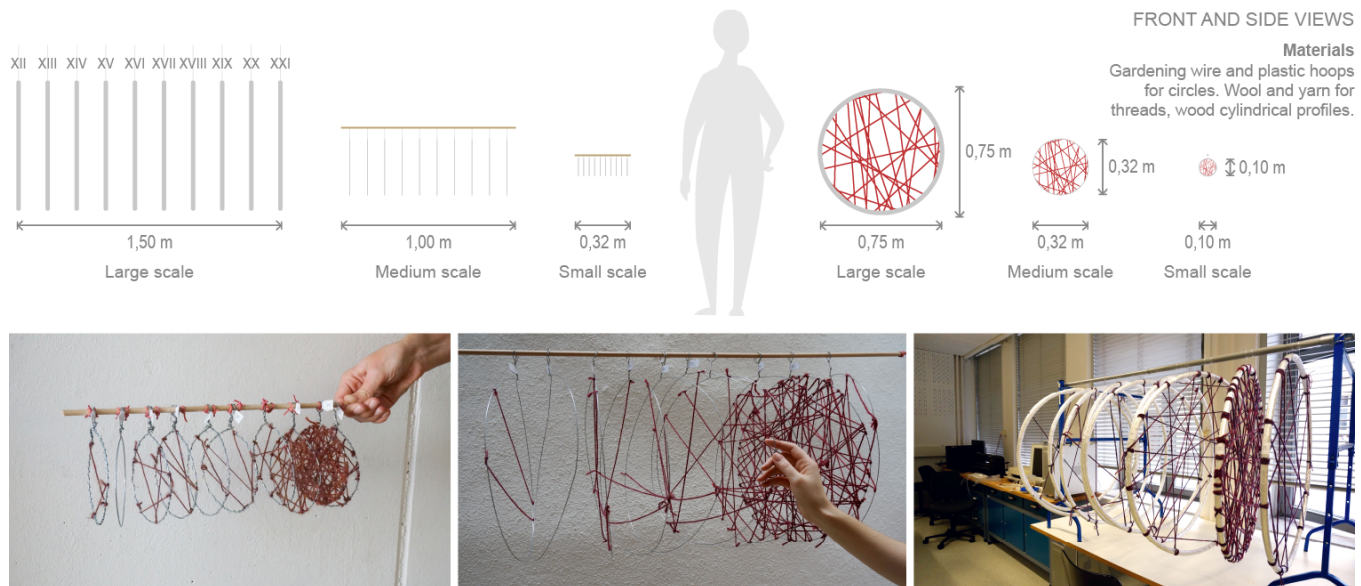

Figure 4: Design B. Bottom row: Photos from smallest (0,32x0,10m) being held by hand over medium (1,00x0,32) to largest (1,50x0,75m) version. Top row: Diagram of all design versions (because of the high 3D nature of the object this shows size horizontally and in depth for the hoops) in relation to a human body.

data by century, resulting in 10 hoops. As can be seen in the photos in figure 4, this results in a representation that is fairly quick to visually grasp. Both designs have a clear mapping that is easy to decipher (evading the problem highlighted by Jansen et al. [25] of 3D volume being hard to compare), encoding the data in the length of thread or the number of threads on a hoop, respectively.

The designs are based on two different metaphors. Design A uses a subtractive approach: "peaceful" white and ordered threads are increasingly cut off. Emptiness here reminds of the reduction of diversity and its irreversibility. In contrast, Design B displays this as additive damage, as more and more chaotically laid red threads (figure 5). These nevertheless are abstract metaphors (instead of something very direct and visceral, such as using animal bones or

a picture of each species) and each design only employs one such metaphor.

Our goal was not to compare the two designs as such, but to investigate the effect of scaling them, and to do this with very different representations with roughly comparable information density (in terms of being similarly easy to 'read'). Using two designs allowed us to compare results, enabling us to better discriminate between reactions related to the design itself and relating to scale. Moreover, we deliberately did not develop an interactive physicalization, as this could detract from the study focus, creating a novelty effect, as well as requiring unnecessary (for our purposes) effort and complexity in building the objects.

Figure 5: Design B (medium) as a summative view, from past to presence, shows more and more threads of extinction.

Table 1: Dimensions for each object: height x width, width referring to frame length, for Design B also frame diameter.

	Design A	Design B
Large	1m x 1,5m	0,75m x 1,5m, Ø 0,75m
Medium	0,50m x 0,84m	0,32m x 1,0, Ø 0,32
Small	0,21m x 0,32m	0,10m x 0,32m, Ø 0,10

3.2 Three Scales per Design

We built three scale versions per design, resulting in six objects (see figure 3 and 4). Both small versions are hand-held, and the medium can also be handled. The largest version was initially planned to be human-size, but given the need for stability and portability in order to evaluate the constructions in public settings, the dimensions were reduced to 1m and smaller height (see table 1); this 'large' version thus corresponds is effectively furniture sized and not of architectural scale.

All objects were built by hand, from crafts material and other purchased materials. For design A, three very similar versions were created. Each scale had thread of different thickness (wool and yarn), roughly scaling the design proportions. For design B, the threads, similar to design A, were from wool and yarn of different thickness. Design B posed construction challenges that required some compromise in scaling. The small version was built first. When building the medium size, the hoop wire began to have difficulty maintaining its (circle) form. This version had a cleaner attachment of threads to the hoop than the small version, which was later noticed and commented on by participants. For the large version, the hoop wire proved too weak to maintain shape. Thus, we purchased more stable plastic hoops to use (these were the closest available to the desired scaled-up size). Unfortunately, compared to the smaller versions, these were much thicker in proportion to the threads, and thus far more visible than the smaller hoops. Thus, the result of scaling was not as satisfactory as for design A, influencing our findings, as we will discuss.

4 USER STUDIES

A between-groups questionnaire study with 85 respondents provided ratings for all 6 objects. For deeper insight, this was augmented with a smaller number of interviews, where interviewees experienced all three versions of a design.

4.1 A Questionnaire Study in Public Spaces

We recruited members of the public at public spaces (figure 6). Each object was placed individually at a location, and 12 to 15 participants were recruited per object as convenience sample. There were six locations in total. We sought out high-traffic areas such as bus stops, the train station, or recreational areas, where passers-by might have spare time. Also, transportability was a factor. Design A was shown mostly in outside public spaces. Design B, which was more voluminous and unwieldy even at medium scale, was mostly shown in closed, indoor spaces related to the University (e.g. the Mensa), resulting in a younger demographic (no respondents aged over 60). The larger structures were placed on the floor, the medium-sized versions were placed on a table (design B) or a low wall (around 0,40m high, design A), and the smaller structures were directly handed to participants.

One researcher set up each object for 2-4 hours at a site. An A4 printed sign "Species extinctions" (in English and German) was put by its side, with a smaller subtitle indicating this was a study for Bauhaus-Universität Weimar. In total, 85 passers-by took part, often effectively self-recruiting by approaching the objects driven by curiosity, in particular for the medium and large scale. For the hand-held size versions, most respondents had to be approached directly. Anybody who then declared willingness to participate in the study, first received an explanation of what the represented data referred to (species extinction over time) and then the questionnaire was introduced. Verbal comments from respondents were noted down in writing, omitting any recording. Once the participant was done, we attempted to recruit the next person.

Each object was rated by 12 to 15 respondents. 49 participants answered the German, and 36 the English version of the questionnaire. Doing so took 5 minutes maximum. We asked only basic demographic information. 42 respondents were female, 42 male and one chose not to identify. The biggest age group (55) were young adults (age 18 to 34), followed by 35-59-year-olds (18 respondents). The first set of questions asked on a 5-point Likert scale (these focus on data understanding and engagement with the topic):

- I understand the represented data (Comprehensible)
- I think the information displayed is interesting
- This makes me feel concerned about the extinction of species (Concerning)
- This makes me want to inform myself better on the subject (Research-worthy)
- This makes me want to take action on the matter (Action-worthiness)

The second part contained the short version of the AttrakDiff questionnaire, (taken from [5]). We selected the AttrakDiff as it is an established UX questionnaire that focuses on a wide range of aspects and assessment criteria in user experience: pragmatic quality (i.e. usability and usefulness), hedonic quality (style, but also emotional needs, such as stimulation and identity) and attractiveness [16, 19].

Figure 6: Some examples of versions on-location, from the local park over a city square to the University canteen/Mensa.
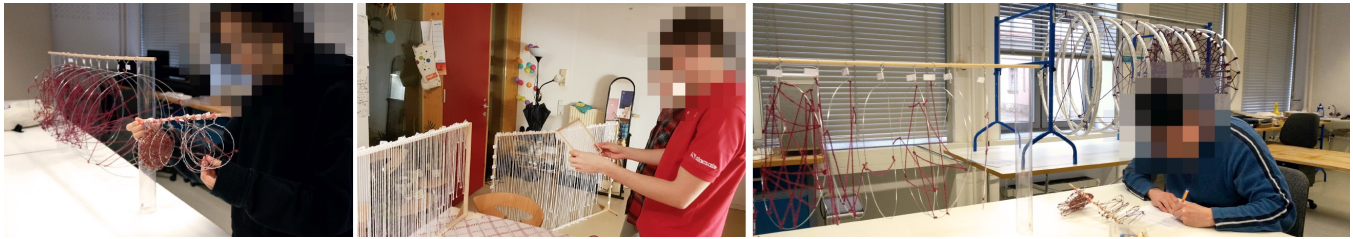

Figure 7: Participants exploring the 3 versions of Design B (left, right) and Design A (middle).

It is a measurement instrument based on semantic differentials (opposing adjectives). Respondents rate the product/object on a range of seven steps for each adjective-pair (such as confusing-clear, unimaginative-creative). While the full version has 28 items, the short version consists of 10, resulting in a shorter questionnaire, which better suits our setup with respondents recruited in-situ in public locations.

The first section of the questionnaire was analyzed based on average scores. For the AttrakDiff, averages were calculated and adjusted to the negative and positive value scale proposed by the AttrakDiff website. We furthermore calculated standard deviations for all items of both questionnaires to determine agreement in ratings (this is normally not done for the AttrakDiff).

Our study is a between-groups study. Moreover, all respondents saw only one version of one design. Our findings are thus based on descriptive statistics which can only indicate trends and are vulnerable to chance effects. Moreover, the location might influence responses (context) and yield different demographics (e.g. University canteen = students). We checked for the latter by analyzing responses according to age group and could not find any systematic differences in responses.

4.2　Semi-structured Interviews

To reveal deeper and richer insight into a viewer's experience of the representations at different scales, we also organized six semi-structured interviews, where participants were invited to investigate the three versions of a design. Two interviews involved pairs (8 interviewees in total). Participants were mostly students, recruited by approaching them personally (4 women and 4 men, aged 22 to 37). Based on our experiences from the questionnaire study, we decided to focus on Design B, which was faster to understand and enabled viewers to focus on trends. Nevertheless, we also ran two

interviews with Design A. Overall, two interviews (3 people) were done with Design A, and four interviews (5 people) with Design B.

A quiet and secluded room was chosen so the interview would be undisturbed. All versions of a design were present in the room, laid around (see figure 7). The largest stood on its own (design A) or on a separate table at eye level (design B) while the medium and small sized were placed on a table next to the chairs of participants. Participants were explained the data represented and instructed of what the study would entail, and then invited to look around and to manipulate the three versions. They then read and signed the consent form. Once they were comfortable, audio-recording was activated. First, we asked participants for their preference for one version (i.e. scale), and then they were interviewed along similar themes as in the first section of the questionnaire, probing into how they felt the three versions differed, e.g. which one was easier to comprehend, more attractive, where they could imagine them to be used, if one made them think about species extinction more than the others, etc. When interviewees had no further thoughts, the interview was ended and the participant thanked. The audio was transcribed and coded for analysis. The interviewer also took notes on the bodily behavior of the interviewees.

Interviews lasted 10 to 20 minutes, and participants were free to roam the room to inspect the objects. Only three people stood up to look closer at the larger piece, while all of them held the smaller version in their hands at some point. The medium-size was often pointed at, but not directly touched during discussions.

5　FINDINGS

Against our original assumption that larger size would have the most impact, we found that overall the medium sized versions of both designs received the most favorable responses. We present results from questionnaires, before reporting interview findings.
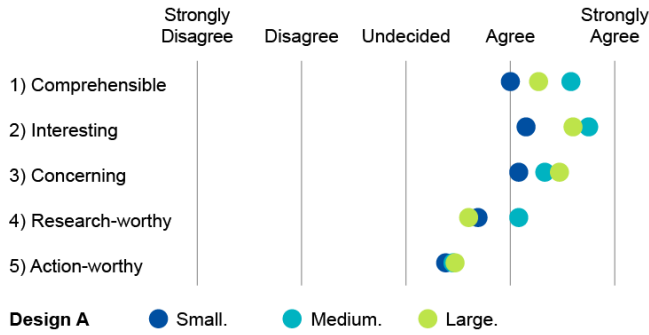
Irene López García and Eva Hornecker

Figure 8: Ratings for Design A: The medium version was rated slightly better than the others apart for 'concerning'. The small version fared worst on almost all criteria.
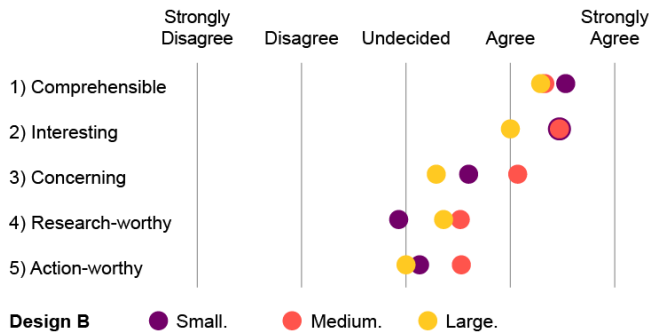

Figure 9: Ratings for Design B: The medium version fared slightly better on all criteria apart for comprehensibility. The large version was rated worst on most criteria.

5.1 Questionnaire Results

Given the comparatively small number of respondents, who each only saw one version of one design, questionnaire results need to be interpreted with caution. In the first part of the questionnaire, the medium scale of both designs fared slightly better than the small and large version on most of the criteria (being understandable/comprehensible, interesting, raising concern about extinction, raising desire to learn more, and impetus for action) (figures 8 + 9). For Design A, the smallest scale (dark blue) fared worst, whereas for Design B, the largest scale (yellow) fared worst. None of the designs made people more inclined to act on species extinction, with responses hovering close to 'undecided'.

Agreement on ratings was highest for the medium version of design A (in order of questionnaire items: SD= 0.67, 0.45, 0.78, 0.79, 0.93), with an average SD over all items of 0.72). Rating agreement for both small versions was slightly lower with an average SD of 0.8 (for A-Small: SD= 0.82, 0.8, 0.76, 0.78, 0.87) (for B-Small: SD= 0.64, 0.52, 0.98, 1.1, 0.74). Agreement regarding the rating of B-Medium and A-Large was lower, but still with an average SD<1 (for B-Medium: SD= 0.97, 0.83, 1.1, 0.64, 0.99, average SD=0.91) (for

A-Large: SD= 0.8, 0.63, 0.83, 1.35, 1.12, average SD=0.95).[1] For B-Large, respondents agreed the least on any of the questionnaire items, with standard deviation for all items >1,0, and an averaged SD=1.27) (in order of questions: 1.27, 1.11, 1.47, 1.08, 1.47).

B-Large thus had both the worst overall rating and the least agreement on ratings. A-Medium had the highest rating of design A versions and the overall highest agreement. Participants also had comparatively good agreement on the low rating for A-Small.

Figure 10 and 11 show the standard AttrakDiff representation for comparing objects.[2] According to AttrakDiff ratings, the large version (light green) of design A fared best, closely followed and occasionally surpassed by the medium size (figure 10). The small version fared worst on all categories (dark blue line), with up to almost two points difference on the 7-point rating scale. For design B (figure 11), the smallest version (dark purple) was rated somewhat better on most criteria, but no pattern was found for the other versions, and overall, differences in ratings are less pronounced as for design A. Overall, A-Small was rated worst of all.

Although not normally done for the AttrakDiff, we calculated Standard Deviations. Whereas the previous part of the questionnaire was on a 5-part Likert scale, this is on a range of 7. For design A, the small version (rated worst) had the highest overall agreement across AttrakDiff categories (averaged St.Dev =1.23; 4 items with SD < 1, another 2 with SD <1.2, and 3 items with SD>1.5), thus there was a tendency for high agreement on a low rating (we can trust the dark blue line for most categories). The medium version also had good agreement (averaged St.Dev =1.33; 3 items with SD < 1, and 3 items with SD>1,5). The large version had the least agreement (averaged St.Dev =1.43, 4 items with SD<1.2, and 5 items with SD>1.5). Both for A-Medium and A-Large, participants tended to disagree on the item cheap-premium (SD=2.06 and 2.09).

For design B, the small and medium version had also good agreement (B-Small: averaged St.Dev =1.25; 3 items with SD<1, and 4 items with SD>1,5) with the medium version having the most consistent ratings (B-Medium: averaged St.Dev =1.21; 3 items SD<=1.1, only 1 item with SD>1.5). B-Large had the most divergence in ratings (averaged St.Dev =1.57; only 2 items with SD <= 1.1., and 7 items with SD>1.5), with the least agreement for practical-impractical and stylish-tacky (SD=1.98 and 1.9). The lack of a clear pattern in ratings for design B shown in figure 11 is thus aggravated by lack of agreement for B-Large (yellow line).

Most participants initially struggled with the concept of semantic scales and were confused about some adjectives, needing further explanation. In particular, pairs such as "predictable – unpredictable" did not seem to apply, being more suited for interactive products. Furthermore, a few participants believed data from Wikipedia might be unreliable, causing skepticism towards the data.

Despite its similarity to standard bar charts, design A overall needed more explanation. Design B, on the other hand, was very quickly understood. Furthermore, with design A, people tended to

---

[1]One participant who rated B-Medium appeared to be a statistical outlier who felt neither concerned about species extinction after seeing the data nor found the information interesting, giving ratings 2 Likert points away from all other respondents – this skews SD given low number of participants. For A-Large, one participant was unconcerned, not wanting to inform themselves better or to take action, in contrast to all others. These people might not care about the topic or, on the contrary, be already well informed so our data representation did not make them (more) concerned.
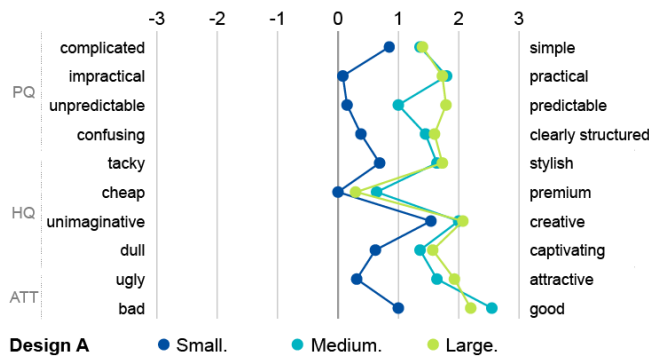[2]Note that AttrakDiff graphs do not show standard deviations.

Figure 10: AttrakDiff Ratings for A-Small are consistently left (worse) of the other lines. The difference between the medium and large version is marginal. (PQ = pragmatic qualities, HQ = hedonic qualities, ATT = attractiveness)
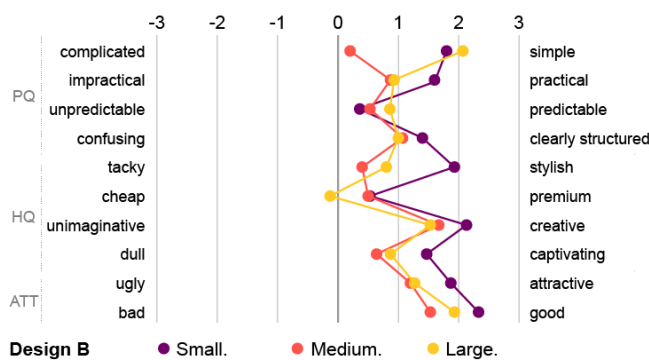

Figure 11: AttrakDiff Ratings for design B. Here the small version was rated somewhat better than the others, but there is no clear pattern.

talk about the raw data, whereas design B also elicited comments on the topic itself and its impact, raising comments and questions, such as, *"What happened here, in the 20th century?"* and *"That seems sad"*. These observations motivated us to focus on design B for the interview study.

5.2 Findings from comparative interviews

Our insights from the interviews, where participants could see all three scale versions of one design and interact with them (see figure 7), concern size preferences, legibility, and emotional impact. Preferences for a scale version were frequently influenced by the perceptual legibility of the representations, resulting from material design properties and from the ability to view and 'take in' the representation at once as well as the ability to navigate the data.

5.2.1 Size, Viewability and Navigability. The large versions of both designs were experienced as 'uncomfortable', *"too big"* (P6). This also meant that the annotations were too far away to read. In particular, participants commented that the large versions could not *"be seen as a whole"*, as one *"needs to walk around"* (P1) and the information is not as *"concentrated"* (P1) or *"compact"* (P6) as

with the other versions. The large size apparently made comparison difficult, as the eyes have to refocus on a data point at a time, then move over to repeat, which results in losing overview: *"(you) cannot compare it in a whole image"* and lose focus (P1). P7 remarked on design A: *"the biggest one, is too big for my eyes. I don't know where to concentrate, because of this scale"*.

With design B, the medium-scale version was mentioned as *"suitable to look at as a whole"* (P1) (or: *"I can see the whole at once"* (P0 on design A)). P4 also discussed the medium version: *"And you can put yourself in like just one point of view and see through all the circles as well"*, emphasizing that this works best for the medium version. Participants implied that it was easier to move around this version to analyze the different perspectives: *"you can go all around […] and you can see like through it"* (P4).

5.2.2 Size Preferences. Participants were asked for their preference for a size version. For both designs, medium size was clearly preferred (by 7 of 8 participants). As discussed above, the medium version of both designs made it easiest to see the representation all at once, and to navigate and view it from all sides. Furthermore, especially for the medium version of design B, referred to as the most 'clear' or 'clean', and described as 'elegant', the relative size of threads and hoops worked best, whereas with the large version, the physical hoops appeared to distract from the data (the threads) (this could be interpreted as figure-ground).

Only P3 preferred the smallest version (of design B), aesthetically and as data representation, and mentioned it had the best proportion between hoop and thread size, talking later about the *"size of the hands"*, while playing with the hoops. Most of the interviewees for design B tended to play with the small version, to inspect or talk about different states and views, making the hoops move and turning the construction around to reveal different perspectives. Other than this, participants did not show much appreciation for either of the small-scale representations. P1 and P5 said it was *"not impressive"* and *"too small"*. P7 found the small version of Design A too dense: *"everything is like vague, and one big mixture. I don't understand what I see."*

When talking about the topic or the data, people kept looking at the medium scale representation, pointing at the hoops or threads. They sometimes also used the small version to do so, but less so. The large version was only referenced in this way when explicitly mentioned (note: this might also be because it was placed on the ground, furthest away, thus being out of physical reach, whereas the other constructions were placed on the table).

5.2.3 When to Use – Large to Impress, Small Hard to Share. Despite preferring the medium size version and finding the large version uncomfortable to view, when talking about which version would be best to impress other people, all suggested the largest scale. Two participants referred to the large scale of design B as 'surprising'. It was imagined as an installation piece or sculpture for public spaces. The smallest version was deemed unimpressive, and often commented on as not shareable for small-group situations (P1: *"too small for everyone to look at"*), albeit it could work as a *"conversation piece"* on a coffee table (P5).

5.2.4 Emotional Impact. All found it difficult to emotionally connect. This could be expected, given the representations were intentionally abstract. One participant noted a potential conflict between design aesthetics and message: *"[…] there's a conflict between a strong, something very sad, very… we should be concerned about, but on the other hand it looks pretty"* (P6). However, P0, P1, P2, P3 and P4 mentioned that the physicalizations made them think about the topic and wonder about a 'jump' in the 20th century, indicating curiosity and concern. This concern appeared to be more intellectual than emotional. P3 said *"It helps me understand what happened (…). But it's not like a shocking effect that it triggers in me, the necessity to do something."* P4 commented on the representation as simple and easy to remember, indicating a potentially lasting impact:

"But I think because it's more simple and informative, it's easy for it to enter your mind. And because it's very simple I think it's easier to memorize. I think I will always have the image of the threads in my head when someone talks about extinct animals (…). Because it's very clear (…) it's easy to memorize, in a way. I can feel that the image is already stuck in my head."

5.2.5 Influence of Material Design . The three version of design B were built with different materials and, as a result, not exactly proportional in scale. This clearly constitutes a limitation of our study, since participants' ratings and preferences were influenced by this. Nevertheless, for future work these are important lessons, as it was not evident from the outset how much this would impact responses. P2 and P3 mentioned differences in roughness and the relation of threads and hoops. P2 added that the thicker hoops of the large version were distracting and *"rough"*. P5 said they were *"in the way of the data"*. With the large version, the hoops are far more prominent, which seemed to both distract from the threads and to impact aesthetics. The small version was perceived as more dense and 'rougher' than the medium version, with the threads being thicker in relation to hoop diameter, and the fastening of threads around the hoop more visible.

It was already mentioned that the movable hoops of design B invited playful handling, especially for the small version. In contrast, those interviewed with design A rarely picked up the construction, and when doing so, acted rather carefully. This could be due to the thin white thread, which makes the object look fragile; in addition, there are no moving pieces that might invite fidgeting and play.

5.3 Summary Overview of Findings

The interviews allowed for a direct comparison of all three size versions and provided rich data, participants giving a clear and consistent response regarding their preferences. The questionnaires do not give a clear picture, likely because participants only saw one size version each, and due to the smallish survey. AttrakDiff ratings only show a clear pattern for design A, where the visual design was scaled more consistently than for design B. While insights from interviews are clearer, the questionnaire study nevertheless was valuable in informing our further approach, as well as providing implications for further research.

Responses to the first part of the questionnaire, which focused on informational value and impact on participants' thinking, and the preferences given in the interview as well as nonverbal references (looking at/ pointing) made during interviews coincide for both medium scale designs, as well as for B-Large and A-Small, but contradict for A-Large (which did well in both questionnaire parts). Both small versions were 2nd ranked in the interviews (neither preferred or disliked, but pointed at, and with B-Small, being 'played' with most, indicating good affordances for interaction). A-Small did worst of all three sizes of design A in both questionnaire parts. It was commented on as visually 'dense' and "vague" in interviews, but otherwise received little explicit negative comments.

AttrakDiff responses do not correspond to interview preferences (very clear for A-Large, B-Small and B-Medium). In particular, A-Large was rated as more practical, clearer and predictable than the small version, as more stylish and attractive, despite being rated as less comprehensible in part 1 of the questionnaire. Interestingly, both A-Large and B-Large had the least agreement in AttrakDiff ratings (while being most disliked in interviews). This might be partially influenced by the large versions being perceived as more impressive (mentioned in interviews). This indicates that the AttrakDiff's focus on hedonic qualities and attractiveness can skew preferences towards impressiveness, if participants see objects on their own, without a comparison.

6 DISCUSSION

6.1 Resizing Material Design may not Scale

While evident in hindsight, it was not clear from the outset how difficult it is to scale every aspect of a physicalization. Digital representations can be scaled virtually, but material design can be hard to scale up/down (a problem noted in prior research on installation design [20], resulting in differences beyond scale, which complicates or disables experimental comparisons (changing more than one variable). This is important for future research on the impact of scale. In our case, scaling required finding new materials, e.g. thicker or thinner string, which then was not perfectly scaled or had different appearance from its fabric, resulting in altered aesthetics or perceived density of strings. Moreover, physical constructions will often require changes in construction or material when resized, to retain stability (in our case, the hoops) because physical forces do not always behave linearly (scaling fallacy). What is stable in a small model might require extra trusses and counterforts for a large version, or different contours, or a change of materials, if the construction method cannot be scaled (e.g. when the human-size version of a 3D print cannot be 3D-printed anymore).

6.2 Which Scale is Best?

The role of size and scale for physical 3D representations has been rarely focused on in HCI. Mirroring earlier findings regarding screen size [40], the ability to see the entire physical representation at once, so that extreme values can be compared at a glance, without 'having to refocus' or turn head or eyes, influences people's preferences. Despite not being huge, and our participants being allowed to physically navigate around the objects, both large versions were commented on as requiring more effort and being uncomfortable, and were therefore not preferred. They were only preferred in case the goal would be to impress and attract attention (possibly an underlying reason why design A-Large did well in the AttrakDiff, which includes criteria on attractiveness and hedonic qualities). Our findings indicate that the size of physicalizations needs to be

well-chosen so as to not interfere with ease of perception (unless this is deliberate).

6.2.1 Advantages of Smaller-Scale: Physical Handling . Besides of being quicker to 'take in', that is, see at once, smaller physicalizations enable the user to physically handle and quickly change perspective by turning the object around. Participants particularly liked handling the small and medium size version of design B, which allowed for manual exploration and possibly appealed to playful interaction through the individually movable hoops.

As discussed earlier (section 2.2.3), research suggests that touching objects supports memory and creates emotional connection [17, 18, 23, 27, 28, 43, 53]. A preference for smaller objects has also been noted by Lee et al. [29] who investigated how the size of mobile handheld devices affects users' gestural interaction and their preferences for a size option, with smaller handheld touchscreens requiring less manual effort. For gestural interaction with screens, kinaesthetic cues may support recall [48]. Pointing at something and then refocusing was easy to do with the small and medium size versions, but the large versions of both of our designs required too much effort and did not provide good overview. The small and medium scale versions thus had the advantages of being easier to see 'at once', requiring less refocusing of the eyes, as well as being easier to manually interact with.

6.2.2 When Might Large Scale Matter? For the reasons stated so far, just scaling up a physicalization does not always work. For large scale representations, effort needs to be taken to ensure it is easily perceivable and supports comparisons or viewers' focus shifts. What distances do the eyes need to accommodate at, what can be taken in at a glance, without turning the head?

This implies that with large-scale physicalizations, their scale should matter and provide extra value – for instance being impressive, attracting attention, or letting viewers experience things in relation to their own body, contributing metaphorical meaning. As existing large-scale physicalizations show, scale can contribute to the interpretation of data. Knowing that in 'Of all the People in the World' [8], each element is just a tiny corn of rice and seeing that the resulting heaps are higher than a human body provides an intuitive, embodied understanding of abstract numbers and magnitude. Smaller grains, resulting in little heaps would not be as striking and direct. In 'Mount Fear' [39], the mountains occlude sight of the map. Similarly, the flow of poppies filling the Tower of London in 'Blood Swept Land' [11] provides an image of being surrounded by a sea of red ceramic poppies. Moreover, visitors walk through the installation, where bodily navigation provides an embodied understanding of distance and space. This creates a far more visceral experience than if one would be looking at a miniature version. Here scale emphasizes and renders abstract numerical quantities concrete by putting them in relation to the viewer's body. Given most people struggle to understand large magnitudes, this is a useful consideration for InfoViz aimed at a general public. Other large scale structures in public spaces, such as Orbacles [30], attract curiosity (aligning with our participant's suggestion of using large representations in museums). Other physicalizations might be optimized for being hand-held and manipulated directly. Many physicalizations of personal data are made to be held in the hands [7, 27, 45], which corresponds to showing personal, intimate data.

In this context, it is interesting that participants were more likely to physically explore the hoops of design B and that this design in the questionnaire study evoked more comments on the topic and its impact, whereas design B had people talk more about the data as such (as numbers).

If the aim of data physicalization is to be a call to action, should the scale which ensures a better understanding of the data be chosen over a visually shocking one? It may well be that in the longer run a more comprehensible model has a bigger impact than an "impressive" one. This is an open question. Given the "performance-preference dissociation" of preference often not being aligned with performance [22, 58], future work should assess the memorability of physicalization at different scale.

These considerations furthermore suggest that scale might not be a parameter we can easily study in experimental studies, because several aspects (aesthetics, viewability, way of interaction) are altered with a change of size, and longer-term effects need to be investigated. When doing such studies, we need to carefully select and design representations that 'work' at different scale and do not lose viewability, or to focus instead on the qualitative experience of a different scale. Comparisons might reveal trade-offs, where one size is quicker and less effort to understand, whereas another attracts more attention and is more memorable, or where one size provides a better general overview whereas the other allows for more active exploration of detail, or a smaller scale supports a more analytic understanding, whereas a large scale might provide an intuitive understanding of magnitudes.

6.3 Limitations and Future Work

Our work has only begun to explore the impact of scale. We deliberately did not develop an interactive physicalization, since this would potentially detract from the focus of the study, creating too much of a novelty effect, as well as introducing further confounds into the generation of three different size versions. There are a number of limitations to our study. First, it had a small number of participants. A larger sample might result in clearer patterns in questionnaire ratings. Having participants rate all scale versions might further contribute to clearer patterns in ratings. In addition, scaling the constructions was trickier than anticipated, resulting in different execution and appearance, especially with design B. Also, the large size for B constrained transport and placement, which biased participant recruitment. Our designs were not made to read exact values, but to allow an overview of the data. This worked better for design B, where values were more freely interpreted by users. To evaluate the readability of values at different scales, more precise designs should be considered.

The rather noisy questionnaire data indicate that future studies should take a comparative approach, letting participants compare and rate all scale versions, in particular since impressive size may skew responses for hedonic, aesthetic and pragmatic (usability) aspects. For interviews, the study design should be adapted, for instance focusing more on the data topic and analyzing which scale representation is used and referred to most when talking about the issue at stake, or used in which way, avoiding explicit questions about preferences, which can suffer from post-hoc rationalization.

Regarding our research question on the impact of scale, neither the large version of design A nor B was really 'architectural', or truly human-sized. Our study thus might not yet fully assess the impact of size. Given a well-chosen design, we should therefore investigate whether a truly LARGE size elicits a more emotional response or deep engagement with the topic than a hand-held object. Furthermore, data representations with a higher metaphorical value might have more emotional impact, which then might be experienced differently at larger scale. Moreover, future work should investigate whether a larger, more 'impressive' scale or a smaller version that allows for hands-on exploration results in increased memorability (and whether this refers to the overall topic and message, or to details of the data).

7 CONCLUSION

With physical data representations, scale/size is clearly an important design factor. Yet the role of scale has rarely been investigated in data physicalization research so far. In particular, there has been no work studying its influence on user experience. Our study reveals that indeed scale influences users' reactions, but that its impact on representation legibility and affordances for action needs to be considered in design. While we have only begun to study the impact of scale, our work highlights avenues for further research as well as important design criteria, in particular the influence of scale on legibility (ability to 'take in' as a whole, combined with being able to attend to details, and to shift between overview and detail). More research on the trade-offs of scale on visibility and navigability is needed. Our work furthermore highlights the challenges and confounds that make it difficult to study the impact of scale for immersive physical data representations.

Firstly, our efforts revealed that it is very hard to accurately scale a physicalization at different levels, in particular if this is hand crafted. While some of this might be familiar to experienced designers, the impact of (thought to be) small changes was surprising. We also argue not to forget that proportional scaling is not always feasible from a construction standpoint, as physical forces do not scale linearly or material construction methods might not scale up.

Secondly, as our findings indicate, different designs (representation choices) are probably best experienced at particular scale – a good physicalization at architectural scale is different from one in handheld size. Thus, design and scale should not be independent decisions. When deciding on scale, the designer should consider the affordances for viewing, manipulation and navigation, which are different at different scale, as well as whether scale contributes to the meaning being conveyed or helps to convey it more easily, providing 'extra value' beyond the wow-factor. If scale is an integral design aspect, then it might in fact be almost impossible to do a comparative study on the impact of scale (since each size version might have different affordances and add more or less to the meaning conveyed, or enable more detailed reading of data items versus an overview of trends) – but it was only by reflecting on the outcomes of our study presented here, that these factors and difficulties were revealed. We hope this reflection and our insights to be useful for practicing designers of data physicalizations as well as for researchers in this field, and to inspire and inform further work on the role of scale and size in physicalization.

REFERENCES

[1] Andre G. Afonso, Ecem Ergin, and Ava Fatah gen. Schieck. 2019. Flowing Bodies: Exploring the Micro and Macro Scales of Bodily Interactions with Urban Media Installations. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) *(DIS '19)*. Association for Computing Machinery, New York, NY, USA, 1183–1193. https://doi.org/10.1145/3322276.3322378

[2] Jason Alexander, Anne Roudaut, Jürgen Steimle, Kasper Hornbæk, Miguel Bruns Alonso, Sean Follmer, and Timothy Merritt. 2018. Grand Challenges in Shape-Changing Interface Research. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173873

[3] Christopher Andrews, Alex Endert, and Chris North. 2010. Space to Think: Large High-Resolution Displays for Sensemaking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 55–64. https://doi.org/10.1145/1753326.1753336

[4] Karl Aspelund. 2014. The principles of design: Proportion and scale. In *Designing: an introduction*. Fairchild Books, Bloomsbury, UK, 108–111.

[5] AttrakDiff. 2019. Assessment. Retrieved March 9, 2019 from https://esurvey.uid.com/project#!login

[6] Robert Ball, Chris North, and Doug A. Bowman. 2007. Move to Improve: Promoting Physical Navigation to Increase User Performance with Large Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA). Association for Computing Machinery, New York, NY, USA, 191–200. https://doi.org/10.1145/1240624.1240656

[7] Stephen Barrass. 2014. Acoustic Sonification of Blood Pressure in the Form of a Singing Bowl. In *SoniHED – Conference on Sonification of Health and Environmental Data*, Vol. 12. York, UK, 6.

[8] Stan's Café. [n.d.]. Of All The People In All The World. Retrieved December 12, 2019 from http://www.stanscafe.co.uk/project-of-all-the-people.html See also http://dataphys.org/list/of-all-the-people-in-all-the-world-stats-with-rice/.

[9] Sandy Claes and Andrew Vande Moere. 2015. The Role of Tangible Interaction in Exploring Information on Public Visualization Displays. In *Proceedings of the 4th International Symposium on Pervasive Displays*. Association for Computing Machinery, New York, NY, USA, 201–207. https://doi.org/10.1145/2757710.2757733

[10] Marcelo Coelho and Jamie Zigelbaum. 2011. Shape-Changing Interfaces. *Personal Ubiquitous Comput.* 15, 2 (Feb. 2011), 161–173. https://doi.org/10.1007/s00779-010-0311-y

[11] Paul Cummins. 2014. Blood Swept Lands and Seas of Red. Retrieved December 13, 2019 from https://www.paulcumminsceramics.com/blood-swept/

[12] Schloss Dagstuhl. 2018. Dagstuhl seminar on Data Physicalization: October 28 - November 2, 2018. Retrieved August 22, 2018 from https://www.dagstuhl.de/en/program/calendar/semhp/?semnr=18441

[13] Pierre Dragicevic, Yvonne Jansen, and Andrew Vande Moere. 2019. Data Physicalization. In *Springer Handbook of Human Computer Interaction*, Jean Vanderonckt (Ed.). Springer.

[14] Nick Dulake and Ian Gwilt. 2017. Flying with data: Openness, forms and understanding. *The Design Journal* 20 (7 2017), S3863–S3872. Issue sup1. https://doi.org/10.1080/14606925.2017.1352889

[15] Sean Follmer, Daniel Leithinger, Alex Olwal, Akimitsu Hogge, and Hiroshi Ishii. 2013. InFORM: Dynamic Physical Affordances and Constraints through Shape and Object Actuation. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) *(UIST '13)*. Association for Computing Machinery, New York, NY, USA, 417–426. https://doi.org/10.1145/2501988.2502032

[16] User Interface Design GmbH. 2013. AttrakDiff. Retrieved December 14, 2019 from http://attrakdiff.de/sience-en.html

[17] Steve Guest, Jean Marc Dessirier, Anahit Mehrabyan, Francis McGlone, Greg Essick, George Gescheider, Anne Fontana, Rui Xiong, Rochelle Ackerley, and Kevin Blot. 2011. The development and validation of sensory and emotional scales of touch perception. *Attention, Perception, and Psychophysics* 73 (2 2011), 531–550. Issue 2. https://doi.org/10.3758/s13414-010-0037-y

[18] Ian Gwilt, Alaster Yoxall, and Koutaro Sano. 2012. Enhancing the understanding of statistical data through the creation of physical objects. In *DS 73-1 Proceedings of the 2nd International Conference on Design Creativity (ICDC), Volume 1* (Glasgow, UK). the Design Society, 117–126. https://www.designsociety.org/publication/32468

[19] Marc Hassenzahl. 2004. The Interplay of Beauty, Goodness, and Usability in Interactive Products. *Human-Computer Interaction* 19 (12 2004), 319–349. Issue 4.

https://doi.org/10.1207/s15327051hci1904_2

[20] William R. Hazlewood, Nick Dalton, Paul Marshall, Yvonne Rogers, and Susanna Hertrich. 2010. Bricolage and Consultation: Addressing New Design Challenges When Building Large-Scale Installations. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems* (Aarhus, Denmark) *(DIS '10)*. Association for Computing Machinery, New York, NY, USA, 380–389. https://doi.org/10.1145/1858171.1858244

[21] Samuel Huron, Sheelagh Carpendale, Alice Thudt, Anthony Tang, and Michael Mauerer. 2014. Constructive Visualization. In *Proceedings of the 2014 SIGCHI Conference on Designing Interactive Systems*. Association for Computing Machinery, New York, NY, USA, 433–442. https://doi.org/10.1145/2598510.2598566

[22] Mikkel R Jakobsen, Yvonne Jansen, Sebastian Boring, and Kasper Hornbæk. 2015. Should I Stay or Should I Go? Selecting Between Touch and Mid-Air Gestures for Large-Display Interaction. In *Human-Computer Interaction – INTERACT 2015*, Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler (Eds.). Springer International Publishing, 455–473. https://doi.org/10.1007/978-3-319-22698-9_31

[23] Yvonne Jansen, Pierre Dragicevic, and Jean-Daniel Fekete. 2013. Evaluating the Efficiency of Physical Visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2593–2602. https://doi.org/10.1145/2470654.2481359

[24] 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 *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3227–3236. https://doi.org/10.1145/2702123.2702180

[25] Yvonne Jansen and Kasper Hornbæk. 2016. A Psychophysical Investigation of Size as a Physical Variable. *IEEE Transactions on Visualization and Computer Graphics* 22 (2016), 479–488. Issue 1. https://doi.org/10.1109/TVCG.2015.2467951

[26] Andrea Jelić, Gaetano Tieri, Federico De Matteis, Fabio Babiloni, and Giovanni Vecchiato. 2016. The enactive approach to architectural experience: A neurophysiological perspective on embodiment, motivation, and affordances. *Frontiers in Psychology* 7 (2016), 481. Issue MAR. https://doi.org/10.3389/fpsyg.2016.00481

[27] Rohit Ashok Khot, Larissa Hjorth, and Florian 'Floyd' Mueller. 2014. Understanding Physical Activity Through 3D Printed Material Artifacts. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 3835–3844. https://doi.org/10.1145/2556288.2557144

[28] Susan J. Lederman and Roberta L. Klatzky. 1987. Hand movements: a window into haptic object recognition. *Cognitive psychology* 19 (7 1987), 342–68. Issue 3. https://doi.org/10.1016/0010-0285(87)90008-9

[29] Sang-su Lee, Youn-kyung Lim, and Kun-Pyo Lee. 2012. Exploring the Effects of Size on Deformable User Interfaces. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services Companion* (San Francisco, California, USA) *(MobileHCI '12)*. Association for Computing Machinery, New York, NY, USA, 89–94. https://doi.org/10.1145/2371664.2371682

[30] Minn_Lab. 2017. Orbacles. Retrieved November 29, 2019 from http://minnlab.squarespace.com/orbacles/

[31] Andrew Vande Moere. 2008. Beyond the Tyranny of the Pixel: Exploring the Physicality of Information Visualization. In *2008 12th International Conference Information Visualisation*. IEEE, London, UK, 469–474. https://doi.org/10.1109/IV.2008.84

[32] Andrew Vande Moere and Sandy Claes. 2015. Towards Recognizing the Tacit Qualities of Physical Data Visualization. Workshop on Exploring the Challenges of Making Data Physical, co-located with CHI'15.

[33] Fatemeh Moradi, Mikael Wiberg, and Mikael Hansson. 2018. Scaling Interaction―: From Small-Scale Interaction to Architectural Scale. *ACM Interactions* 25, 6 (Oct. 2018), 90–92. https://doi.org/10.1145/3274574

[34] Ken Nakagaki, Luke Vink, Jared Counts, Daniel Windham, Daniel Leithinger, Sean Follmer, and Hiroshi Ishii. 2016. Materiable: Rendering Dynamic Material Properties in Response to Direct Physical Touch with Shape Changing Interfaces. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 2764–2772. https://doi.org/10.1145/2858036.2858104

[35] Juhani Pallasmaa. 2005. *The Eyes of the Skin: Architecture and the Senses* (2nd. ed. ed.). John Wiley & Sons, Chichester, UK.

[36] Stuart L Pimm, Gareth J Russell, John L Gittleman, and Thomas M Brooks. 1995. The Future of Biodiversity. *Science* 269 (7 1995), 347 LP – 350. Issue 5222. https://doi.org/10.1126/science.269.5222.347

[37] Biljana Puric. 2016. How Scale in Art Influences the Viewing Experience. Retrieved December 12, 2019 from https://www.widewalls.ch/scale-in-art/

[38] Majken K. Rasmussen, Esben W. Pedersen, Marianne G. Petersen, and Kasper Hornbæk. 2012. Shape-Changing Interfaces: A Review of the Design Space and Open Research Questions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for

Computing Machinery, New York, NY, USA, 735–744. https://doi.org/10.1145/2207676.2207781

[39] Abigail Reynolds. 2003. Mount Fear East London. Retrieved December 13, 2019 from http://www.abigailreynolds.com/works/44/mount-fear-east-london/

[40] Mikkel Rønne Jakobsen and Kasper Hornbæk. 2011. Sizing up Visualizations: Effects of Display Size in Focus+context, Overview+detail, and Zooming Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1451–1460. https://doi.org/10.1145/1978942.1979156

[41] Adrien Segal. [n.d.]. Grewingk Glacier. Retrieved December 12, 2019 from https://www.adriensegal.com/grewingk-glacier/

[42] Orit Shaer and Eva Hornecker. 2009. Tangible User Interfaces: Past, Present, and Future Directions. *Foundations and Trends in Human-Computer Interaction* 3 (1 2009), 1–137. Issue 1-2. https://doi.org/10.1561/1100000026

[43] Simon Stusak, Moritz Hobe, and Andreas Butz. 2016. If Your Mind Can Grasp It, Your Hands Will Help. In *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction* (Eindhoven, Netherlands) *(TEI '16)*. Association for Computing Machinery, New York, NY, USA, 92–99. https://doi.org/10.1145/2839462.2839476

[44] Simon Stusak, Jeannette Schwarz, and Andreas Butz. 2015. Evaluating the Memorability of Physical Visualizations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 3247–3250. https://doi.org/10.1145/2702123.2702248

[45] Simon Stusak, Aurélien Tabard, Franziska Sauka, Rohit Ashok Khot, and Andreas Butz. 2014. Activity Sculptures: Exploring the Impact of Physical Visualizations on Running Activity. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (Dec 2014), 2201–2210. https://doi.org/10.1109/TVCG.2014.2352953

[46] Faisal Taher, John Hardy, Abhijit Karnik, Christian Weichel, Yvonne Jansen, Kasper Hornbæk, and Jason Alexander. 2015. Exploring Interactions with Physically Dynamic Bar Charts. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 3237–3246. https://doi.org/10.1145/2702123.2702604

[47] Desney S Tan, Darren Gergle, Peter Scupelli, and Randy Pausch. 2003. With Similar Visual Angles, Larger Displays Improve Spatial Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 217–224. https://doi.org/10.1145/642611.642650

[48] Desney S Tan, Randy Pausch, Randy Pausch, Randy Pausch, Jeanine K Stefanucci, Dennis R Proffitt, and Dennis R Proffitt. 2002. Kinesthetic Cues Aid Spatial Memory. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 806–807. https://doi.org/10.1145/506443.506607

[49] Christopher Tweed. 2000. A phenomenological framework for describing architectural experience. In *Phenomenology and Culture*. University College Cork, Cork, Ireland, 1–9. http://orca.cf.ac.uk/95941/

[50] Björn van der Hoort and H Henrik Ehrsson. 2014. Body ownership affects visual perception of object size by rescaling the visual representation of external space. *Attention, Perception, & Psychophysics* 76 (2014), 1414–1428. Issue 5. https://doi.org/10.3758/s13414-014-0664-9

[51] Björn van der Hoort, Arvid Guterstam, and H. Henrik Ehrsson. 2011. Being Barbie: The Size of One's Own Body Determines the Perceived Size of the World. *PLoS ONE* 6 (5 2011), e20195. Issue 5. https://doi.org/10.1371/journal.pone.0020195

[52] Yun Wang, Xiaojuan Ma, Qiong Luo, and Huamin Qu. 2016. Data Edibilization: Representing Data with Food. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 409–422. https://doi.org/10.1145/2851581.2892570

[53] Yun Wang, Adrien Segal, Roberta L. Klatzky, Daniel F. Keefe, Petra Isenberg, Jörn Hurtienne, Eva Hornecker, Tim Dwyer, and Stephen Barrass. 2019. An Emotional Response to the Value of Visualization. *IEEE Computer Graphics and Applications* 39 (2019), 8–17. Issue 5. https://doi.org/10.1109/MCG.2019.2923483

[54] Data Physicalization Wiki. [n.d.]. Data Physicalization. Retrieved August 22, 2018 from http://dataphys.org/

[55] Wikipedia. [n.d.]. Timeline of extinctions in the Holocene. Retrieved March 19, 2019 from https://en.wikipedia.org/wiki/Timeline_of_extinctions_in_the_Holocene

[56] Wesley Willett, Yvonne Jansen, and Pierre Dragicevic. 2017. Embedded Data Representations. *IEEE Transactions on Visualization and Computer Graphics* 23 (2017), 461–470. Issue 1. https://doi.org/10.1109/TVCG.2016.2598608

[57] Nathan Yau. 2013. *Data Points: Visualization That Means Something*. Wiley.

[58] Johannes Zagermann, Ulrike Pfeil, Daniel Fink, Philipp von Bauer, and Harald Reiterer. 2017. Memory in Motion: The Influence of Gesture- and Touch-Based Input Modalities on Spatial Memory. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1899–1910. https://doi.org/10.1145/3025453.3026001

[59] Jack Zhao and Andrew Vande Moere. 2008. Embodiment in Data Sculpture: A Model of the Physical Visualization of Information. In *Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts*. Association for Computing Machinery, New York, NY, USA, 343–350. https: