City Research Online

City, University of London Institutional Repository

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Structure and Empathy in Visual Data Storytelling: Evaluating their Influence on Attitude

J. Liem[1] , C. Perin[2] and J. Wood[1]

[1]City, University of London, United Kingdom
[2]University of Victoria, Victoria, British Columbia, Canada

Abstract
In the visualization community, it is often assumed that visual data storytelling increases memorability and engagement, making it more effective at communicating information. However, many assumptions about the efficacy of storytelling in visualization lack empirical evaluation. Contributing to an emerging body of work, we study whether selected techniques commonly used in visual data storytelling influence people's attitudes towards immigration. We compare (a) personal visual narratives designed to generate empathy; (b) structured visual narratives of aggregates of people; and (c) an exploratory visualization without narrative acting as a control condition. We conducted two crowdsourced between-subject studies comparing the three conditions, each with 300 participants. To assess the differences in attitudes between conditions, we adopted established scales from the social sciences used in the European Social Survey (ESS). Although we found some differences between conditions, the effects on people's attitudes are smaller than we expected. Our findings suggest that we need to be more careful when it comes to our expectations about the effects visual data storytelling can have on attitudes. Additional material: *https://flowstory.github.io/attitudes/.*

CCS Concepts
• *Human-centered computing → Empirical studies in visualization;*

1. Introduction

Visual data storytelling is receiving a growing interest across academic fields a n d i n dustries (e .g., [S H10, L R IC15, Few17, Den16, Sch17, RHDC18]). It now has dedicated conferences (e.g., Tapestry [tap]) and conference sessions (at CHI [chi], Info-Vis [vis], OpenVisConf [ovc]), competitions (e.g., PacificVis Visual Data Storytelling Contest [BLV*17], Information is Beautiful Awards [iib]), specialized departments in newsrooms (e.g., NYT [nyt], NZZ [nzz]), and publications targeting the business world (e.g., [Dua10, NK15]).

Visual data storytelling is often said to make visualizations compelling, memorable, understandable, engaging, or persuasive (e.g., [MLF*12, Fig14a]). However, from an empirical perspective, there is little evidence supporting such claims. Only a handful of studies have investigated these questions through controlled experiments that looked at the effect of visual data storytelling on memorability [HDR*13, BVB*13, BBK*16], at persuasive visualization [PMN*14, Mue12], at data anthropomorphizing [BPE*17], and at the relationship between storytelling and level of user-activity [Dia10, BDF15] or engagement [MRL*17]. No study has looked at the effect of visual-narrative techniques – the tools used in visual data storytelling to create narratives – on people's attitude.

Contributing to this growing body of work, we investigate to what extent selected visual-narrative techniques often used in visual data storytelling influence people's attitudes towards the contentious topic of immigration. We conducted two crowdsourced between-subject experiments with 300 participants each to compare (a) personal visual narratives designed to generate empathy; (b) structured visual narratives of aggregates of people; and (c) a fully exploratory visualization without the use of any visual-narrative techniques, which served as a baseline. We ran two experiments to account for the fact that prompting participants with questions about immigration before they were exposed to the stimulus might bias the results: in *experiment 1*, participants were asked questions about immigration both before and after the stimulus (pre-post-test design), while in *experiment 2* they were asked these questions after the stimulus only (post-test design). This single difference between both experiments highlights the difficulty in evaluating the effects of storytelling: experiment 1 yields more power by comparing values before and after the stimulus; while experiment 2 avoids priming participants by not asking questions before the stimulus. We discuss both experiments for reasons of transparency and methodological rigor. We measured attitudes towards immigration using well-established scales from social sciences used by the European Social Survey (ESS) [JRFE07, essb]. While we did

observe statistically significant effects, the magnitudes of the effect sizes are small. Contrary to what we expected, neither personal narratives designed to generate empathy nor structured narrative designed to help people understand facts influenced participants' attitudes strongly.

With this research, we contribute to improving our overall understanding of visual data storytelling. Findings, which do not support common assumptions, are important contributions, as they prevent us from relying on plausible but unchallenged expectations. This can lead to further work on the mechanisms behind how visual-narrative techniques influence attitudes. We further discuss the limitations of the methods commonly used to evaluate visual-narrative techniques in particular, and storytelling in general.

2. Background And Related Work

Over the past two decades, there has been a growing interest in *visual data storytelling*. Research on the topic includes high-level process models (e.g., [CM11, MLF*12, LRIC15]), design space studies (e.g., [SH10, DKN11, BLB*17, SLRS16, BRCP17]), descriptive frameworks and taxonomies (e.g., [HD11, nap, Rot]), implementations (e.g., [ARL*17, tim]) mainly describing the data storytelling authoring perspective. Little prior research describes the audience perspective, including qualitative and quantitative evaluations (e.g., [HDR*13, Fig14a, BDF15, MRL*17]).

Many contributions focus on the use and evaluation of visual elements and techniques such as pan, zoom, superimposition, widgets for interactions, navigational devices (e.g., stepper or stroller), and their practicability for visual data storytelling (e.g., [GP01, BLB*17, SLRS16, MRL*17]). Complementary approaches evaluate the applicability of theories and techniques from classical narrative disciplines such as literature, comics and film to visual data storytelling (e.g., [HD11, Fig14b, AHRL*15, BRCP17, Kos17]). In this paper, we follow the latter approach to assess the effect of two established narrative techniques (structure and empathy) in a visualization. Through this approach, we aim to reveal insights about the mechanisms of storytelling as a whole, thus expanding the space of visual data storytelling. We know that narrative techniques help create a story in a reader's mind [HJR10, Rya04]. While this is particularly true in the context of literature and film, this also applies to other media [Rya04] such as still and moving pictures, music, and digital media – which includes visual data storytelling. Still, Empirical studies are needed to assess the conjectural benefits of narrative techniques in the context of visual data storytelling.

2.1. Related Controlled Studies in Visualization

We now discuss previous studies that have assessed the effects of **structure** or **empathy**; or that are related to **attitude** in the context of visual data storytelling.

Visual-Narrative Techniques for Structure. A first group of studies have looked at the *overall structure and sequencing* of visual narratives, showing that: delivery mode (author-driven vs. reader-driven, [SH10]) has a small effect on participants' story reconstruction [BW14]; that people create and prefer hierarchical structures for sequencing based primarily on space, and less on time

or measure [HKL17]; that sequences with parallelism – repeating patterns of transition types – as a structural device are beneficial for memorability [HDR*13]; that visualizations (compared to just text) and animated transitions (triggered through stepper buttons and scrolling) improve reader-perceived engagement [MRL*17]; and that the ability to re-watch and compare visual narratives based on data selections (e.g., location) facilitates understanding [CRS*20].

A second group of studies have investigated the effect of *providing structure through adding narrative sections* to common visualizations, showing that: adding visual data storytelling before exposing people to a visualization did not increase their activity levels and immersion [BDF15]; that adding backstory narratives to crowdsourcing evaluations of visualization tools does not result in higher accuracy and attention and does not result in higher confidence, enjoyability, perceived easiness and usefulness [DBD17]; and that asking quiz questions to structure an exploratory stimulus encouraged people to interact more [Dia10].

A third group of studies have looked at the effect of *visual structure*, showing that: the use of human-recognizable objects such as pictograms, may enhance memorability, comprehension, and recall [BVB*13, BBK*16]; and that animated pictograms lead to higher viewer engagement than static standard charts, which lead to higher viewer engagement than static pictograms and animated standard charts [ARL*18].

Visual-Narrative Techniques for Empathy. To date, only Boy et al. [BPE*17] and Concannon et al. [CRS*20] have investigated empathy. Boy et al.'s experiments comparing anthropomorphized data to standard charts embedded in a narrative context did not reveal conclusive evidence that emotion-evoking narratives have an effect on people's empathy. Similarly, Concannon et al. found that the use of personalised data, dependent on a participant's location, did not lead to more empathy towards the narrative's character compared to data unrelated to a participant's location.

Studies related to Attitude. Studies in persuasive visualization [PMN*14] and cartography [Mue12] showed that visualization types (e.g., bar charts, line charts, and tables) or map styles (e.g., propagandist, authoritative, and sensational) can influence the visually conveyed message. Using persuasion theory models, attitude and attitude change have been used as proxies to measure persuasiveness [PMN*14, O'K16]. Kong et al. [KLK18] further showed that persuasively worded titles of visualizations on controversial topics can influence the perceived message but do not have a meaningful effect on attitude change. A first study investigating persuasive data videos indicated that such videos have the *potential* to influence peoples' attitude [CSF*19]. Heyer et al. [HRR20] found that the elicitation and incorporation of participants' prior knowledge in a visual narrative did not lead to significant effect on attitude change. However, they showed that visualizations are generally more persuasive compared to textual representations.

Study Designs. The aforementioned studies test either (a) *objective performance measures* including accuracy, correctness, recall, activeness or time (e.g., [BVB*13, BBK*16, BDF15]), or (b) *subjective metrics* based on self-reported perceptions of engagement, enjoyability, preferences or comprehension (e.g., [Fig14a, MRL*17, ARL*18, HKL17]), or (c) a combination of them (e.g.,
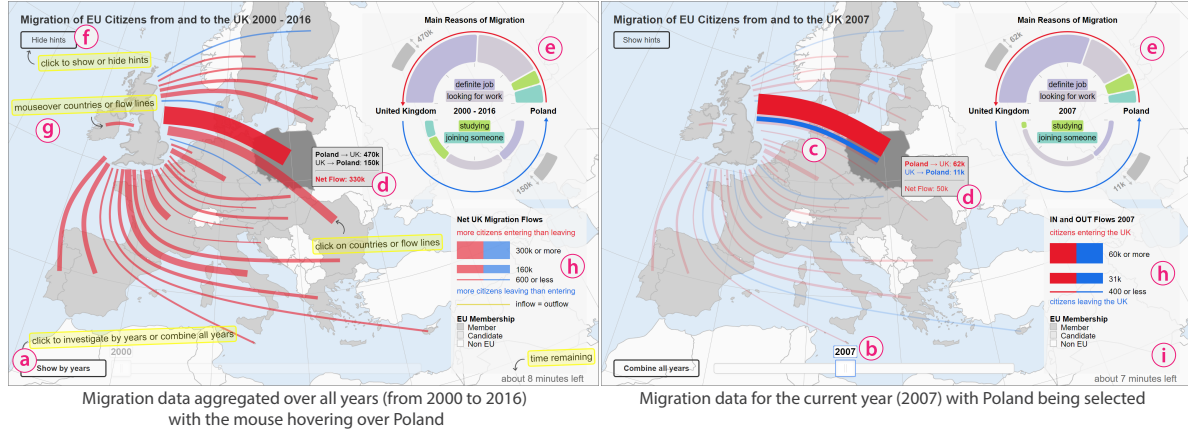
Migration data aggregated over all years (from 2000 to 2016) with the mouse hovering over Poland

Migration data for the current year (2007) with Poland being selected

Figure 1: *Interactive flow map. On the left is the default view. A radial flow map shows aggregated net migration flows (outflow minus inflow) between the UK and other EU countries for 2000 to 2016. A red line means more people are arriving than leaving the UK and a blue line means more people are leaving than arriving. Line thickness encodes net migration flow volume. Clicking on* **button (a)** *switches between the aggregated view on the left and the detailed view on the right. The detailed view shows the data for a single year, that can be set using a* **time slider (b)**. *Here, Poland has been selected (by clicking on the country or on its corresponding flow line). This replaces the net migration flow between the UK and Poland by the inflow (in red) and outflow (in blue) for that country* **(c)**. *Animations reinforce direction of movement, already encoded with red and blue colors. Hovering over lines and countries shows an* **annotation box (d)** *with migration numbers. The donut-like chart* **(e)** *shows the distribution of the four major reasons for migration for the selected country, or for all countries if no country is selected.* **Button (f)** *shows and hides* **hints (g)** *on how to use the interactive map. The* **legend (h)** *provides information about how to read the map elements. A text-based* **countdown (i)** *shows remaining exploration time.*

[DBD17, HDR*13]). Such studies often use between-subject designs (e.g., [BDF15, Dia10, Mue12]) that imitate realistic settings. Following this approach, our study contains a single stimulus that allows to control for learning effects when assessing attitudes.

Previous studies tend to report relatively small effects, often contrary to expectations (e.g., [BDF15, DBD17, BPE*17, BW14, CRS*20]), like in other fields i nvestigating n arrative persuasion (e.g., psychology, advertisement and health) [O'K16, DS12]. This calls for additional studies investigating new effects and conditions in visual data storytelling. The influence of v isual-narrative techniques on people's attitudes has not been studied yet, although attitude has been used as one of the measures to determine persuasiveness [PMN*14, Mue12]). We explain in the next section how people's attitudes are subject to variation when exposed to stimuli, making attitudes a promising dimension to study.

2.2. Values, Attitudes, and the European Social Survey

To explain why and how we looked at the effect visual data storytelling can have on attitude towards a sensitive topic, and why we looked at immigration in particular, we first describe *human values* and *attitude*. **Human values** are *general principles in life, or basic broadly immutable beliefs, that inform more mutable attitudes and opinions* [DM12, O'K16]. **Attitude** refers to *an individual's evaluative judgment about a stimulus object, which can be anything that is liked or disliked (e.g., individuals, groups, events, products and abstract concepts)* [MH15]. While values are quite stable over time, attitudes can be influenced in various contexts and situations [DM12]. This makes attitude a good measure to study, given that it is more influenceable. Human values and attitudes have been collected and analyzed by the European Social Survey (ESS) across Europe since 2002 [JRFE07, essb]. The ESS is a face-to-face survey measuring attitudes towards various aspects of our daily lives (e.g., political engagement, moral and social values, well-being and security). The ESS is relevant to our work because it measures both human values and attitudes towards immigration.

The ESS uses 21 questions to measure ten values from the human value scale [Sch07] (e.g., universalism, benevolence, tradition and security), which can be aggregated into two bipolar dimensions (from *self-transcendence* to *self-enhancement*, and from *openness to change* to *conservation*) [Sch07, DM12, essb]. To measure attitudes towards immigration, the ESS contains seven questions [essa, essb] that identify respondents' opposition to different groups of migrants [DM12], as well as the extent to which they perceive immigration as a threat [ML13]. Research in social sciences [DMBS08, DM12, DMSS14] used the ESS data to show that attitudes towards immigration can be explained by some of these human value dimensions, along with demographic and contextual variables. Previous work found that self-transcendent individuals have a more positive attitude towards immigration, while more conservative people have a more negative attitude towards immigration [DM12, DMSS14].

Our work is based on the assumption that in addition to the relation between less mutable values and more mutable attitudes, visualization – especially the ability visual narratives have to engage with a topic – can play a role in shaping attitude.
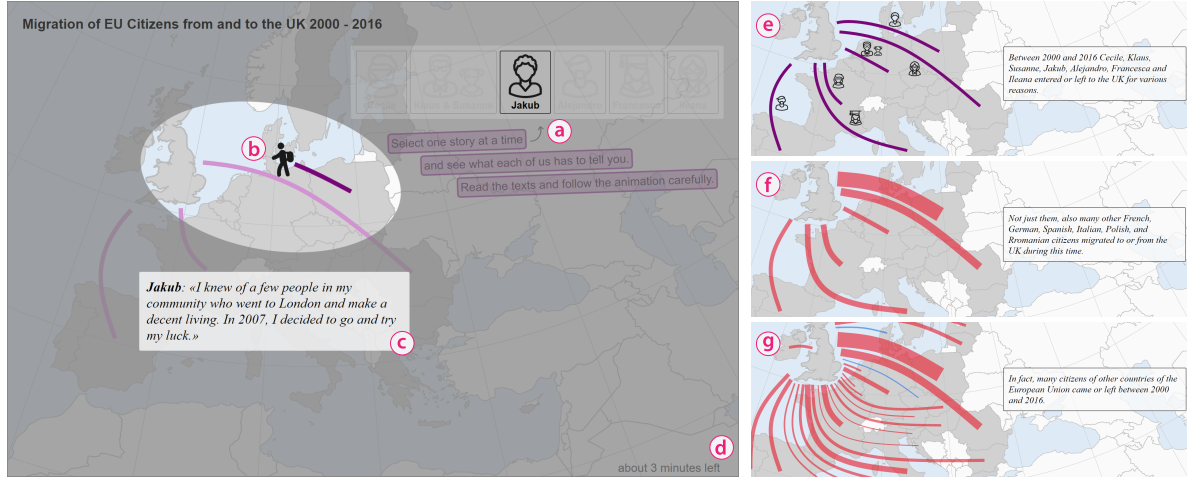
Figure 2: *Empathy design. Selecting one of the six **personal narratives (a)**, plays back the narrative on the map using **animated flow lines and isotype-like characters (b)**. In a **text box (c)**, the character tells her or his personal migration story from a first-person perspective. When the text-based **countdown (d)** in the lower right corner expires, a **transition sequence (e–g)** transforms the view into the interactive flow map.*

3. Experiment Rationale and Materials

To investigate the effect of visual-narrative techniques used in visual data storytelling on attitude, we selected an immigration dataset. Selecting a topic to study changes in attitudes is challenging. We chose a topic that it is close to life and for which people tend to have strongly held attitudes, with the aim of obtaining more meaningful results than if we were to use a topic for which people tend to have weak attitudes. Indeed, while the latter category of topics would likely lead to a larger variance in observed attitude changes, it would also yield less interesting results and open up to the criticism of testing the obvious. The dataset [ons] we used offers enough detail to create meaningful visual narratives (e.g., the reason for migration). We designed an **interactive flow map** (see Figure 1) to explore the data and generate narratives for use in two other designs. We created diverging visual narratives with respect to two categories of visual-narrative techniques (*empathy* and *structure*) to isolate the effect of each technique. For example, while the **empathy** design (see Figure 2) has a first-person perspective and an unstructured story flow, the **structure** design (see Figure 3) uses a third-person perspective and a structured story flow.

3.1. Baseline Design: *Interactive Flow Map*

Figure 1 shows a radial flow map of migration data between the UK and other EU countries. Interactions allow to show the data by year or aggregated, obtain details about specific countries, see inflow and outflow migration volumes, and see the primary reasons for migration to and from the UK. Based on design principles for origin-destination flow maps [JSM*16], we avoided line intersections, used symmetrically curved line shapes, and used an angular distribution around the UK. We used geospatial flow maps because they are a familiar and suitable way to display migration data, and the concepts of space and time are essential in movement visualization and narration.

3.2. Visual-Narrative Design 1: *Empathy*

In the Empathy design, clicking on one of the six available fictional personas makes them tell their personal migration journey, using direct speech from a first-person perspective. Each story unfolds in a chronological order using text. The reason of migration, the act of migrating, and the situation in the UK are mirrored in a visual form using anthropomorphized data graphics [BDF15] (e.g., an unhappy person sitting with the head hanging down). Two of the six personas go back to their home country in their stories to reflect the fact that migrations happen in both directions. Following the idea of the NAPA Card 'humans behind the dot' [RHDC18] we use the six narratives of migration to show that individuals with personal stories are the foundation of the data. This technique is used to *"make abstract data more relatable, and possibly establish an emotional connection between the viewer and the fate of the entities"* [nap, RHDC18]. We also implemented visual hooks through abstract faces and names common in each persona's country of origin (see Figure 2(a)), based on ideas and findings from memorability research [BVB*13, BBK*16]. Finally, we used "story focus," a technique used in narrative maps [MF17] to draw attention to a specific area of the display (see ellipse in Figure 2). A 30-second-long, non-interactive transition transforms the empathy design into the interactive flow map, morphing the flow lines of the single personas to the aggregated net flow lines by country (see Figure 2(e–g)). While each personal narrative possesses a narrative structure, the design is overall unstructured: viewers decide the order in which they want to see the stories. Narrative sequencing techniques are the focus in the next design.

3.3. Visual-Narrative Design 2: *Structure*

The Structure design illustrated in Figure 3 differs from the empathy design in that it presents a structured narrative, uses a third-person narration style, and shows aggregated groups of people (cit-
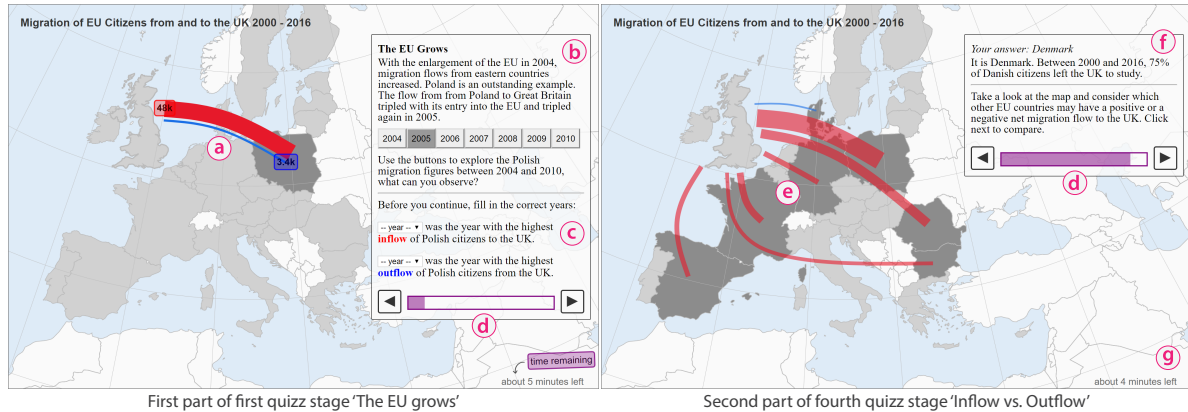
First part of first quizz stage 'The EU grows'   Second part of fourth quizz stage 'Inflow vs. Outflow'

Figure 3: *Structure design: The image on the **left** shows the first part of the first quiz stage of the structure design, showing the **inflow and outflow (a)** of Poland. Labels show the numeric values at the end of the flow lines and indicate direction. The panel on the right-hand side contains: a **text-based narration (b)**; interactive elements to update the map (buttons to select a year in that page); and a **question (c)** that the interactive elements in the panel can help answer. The **navigation element (d)** that resembles a progress bar allows to go forward and backward in the narration. The image on the **right** shows the second part of the last quiz stage, showing **net migration flows (e)** for selected countries. Each second part of a quiz stage provides the **answer (f)** to the question asked in the first part. When the text-based **countdown (g)** in the lower right corner expires, a **transition sequence** transforms the view into the interactive flow map.*

izens of countries) rather than individuals. The structure design consists of six thematically-ordered stages. The first stage provides background information about the data and the map. The following four stages (or 'quiz' stages) each discuss a theme based on one or several countries: The first two discuss the growth of the EU; the third how several countries have had stable migration for a long time; and the fourth the differences between positive and negative net flows. Each quiz stage first shows a narration that introduces the topic, then asks a question related to this topic before the solution is revealed. The last stage is a transition to the interactive flow map, similar to the one for the empathy design. Using repetitions of transitions (question and answer) was found to be beneficial for memory [HDR*13]. Besides providing structure through narrative sequencing, these questions ensured that participants would actively engage with the data space [Dia10] and avoid them clicking through the story quickly. Because stepper and scroller stories with animated transitions increase the level of engagement and immersion [MRL*17], we gave the viewer control over reading speed and the ability to go back in the narrative with the progress bar shown in Figure 3(d). This design was also informed by the NAPA card 'gradual visual reveal', a technique that helps *"the story unfold in the viewer's mind while reading the graphic, to chunk the material, and to make it easier to absorb"* [nap, RHDC18].

4. Studying the Effects of Structure and Empathy on Attitude

To test whether the Structure and Empathy visual-narrative techniques have an effect on attitude towards immigration, we designed two experiments with three conditions each. Both experiments are identical except in experiment 1 we used a pre-post-test design to assess attitudes towards immigration before and after the stimulus; while in experiment 2 we used a post-test design to assess

attitudes towards immigration after the stimulus only (see Figure 4). The design of experiment 1 makes it possible to measure change in attitudes, thus yields more experimental power. The design of experiment 2 yields less power but accounts for conservatism bias [Nis94, ORS09]. Next, we describe the conditions, the study procedure, the data we collected, and the participants. Then we present our hypotheses and report the results. We provide supplemental material, the questionnaire and the interactive conditions online at https://flowstory.github.io/attitudes/.

4.1. Experimental Conditions

The three conditions are based on the designs presented in Section 3. In the **Exploration** condition, participants explored the interactive flow map for 8 minutes. In the **Empathy** condition, they interacted with the empathy design for 4.5 minutes, then the 30-second-long animation transitioned to the baseline design, which they could explore for 3 minutes. In the **Structure** condition, they interacted with the structure design for 5 minutes, then a 2-second-long animation transitioned to the baseline design, which they could explore for 3 minutes. These timings ensured that participants in each condition were exposed to the data for the same amount of time (8 minutes). To ensure that they were also exposed to the same level of detail of information, both storytelling conditions transitioned to the interactive flow map after 5 minutes. The stories in the storytelling conditions focus on the same countries; and they reflect the major reasons for migration in a balanced way.

Before deploying the study, we ran a pilot study with 13 experts in visualization and cartography. This helped us adjust the timings so that the experiment was neither too short (and frustrating for the participants) or too long (and participants would likely lose focus). Qualitative feedback also helped us clarify the instructions.
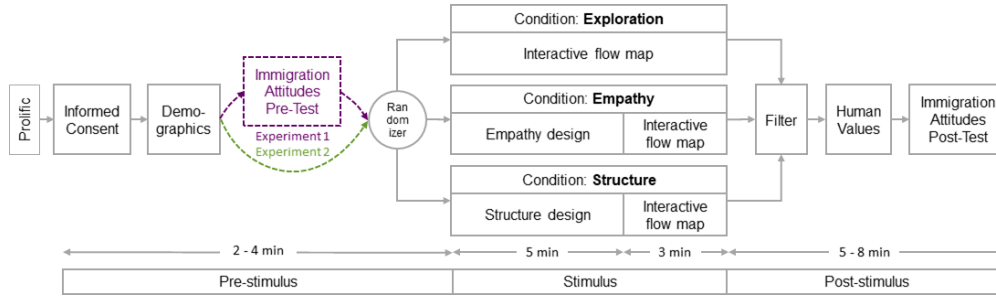
Figure 4: *Study procedures: participant recruitment through Prolific; informed consent and instructions; demographic questions; attitudes towards immigration questions (for experiment 1); stimuli; filter questions; questions about human values and attitudes towards immigration.*

4.2. Questionnaire Design

We designed 28 questions to measure participants' human values and attitudes towards immigration, as described in Section 2.2. All questions had a "Prefer not to say" option, but we asked participants to only use that option if they were uncomfortable answering the question. Because of the sensitivity of the topic, they were reminded that all information provided was anonymous.

The first 21 questions assess human values, from which we can derive two bipolar dimensions: one from self-transcendence to self-enhancement; one from openness to change to conservation [DM12]. The questions are statements describing a person, such as, "She thinks it is important that every person in the world be treated equally. She believes everyone should have equal opportunities in life,". Participants then indicated to what extent they are similar to the described person on a six point scale ranging from "not like me at all" (1) to "very much like me" (6). We adapted the gender of the person in the statement to the participant gender from their demographics questionnaire (or female if a participant did not disclose their gender).

The last seven questions are from the UK version of the ESS [essb] and assess attitudes on immigration. Four questions assess *opposition to immigration* by asking how many people of certain groups should be allowed to come to the UK on a four point scale ranging from "allow many" (1) to "allow none" (4). The groups are people "of same ethnicity," "of different ethnicity," "from poorer countries outside Europe," and "from poorer countries inside Europe." The other three questions assess *perceived immigration threats* on an 11 point scale ranging from "no threat" (0) to "high threat" (10). These questions ask whether immigration is good or bad for the economy, if it enriches or undermines cultural life, and if it makes a place better or worse to live in.

4.3. Participants

For each experiment we recruited 300 participants through Prolific [pro], a UK based crowdsourcing platform focusing on academic studies. Using Prolific's pre-screening tool, we constrained the participation to adults (*age* 18 or older) who are UK *citizens*, were *born* and currently *reside* in the UK, and whose first language is English. These *UK criteria* were used because we selected the topic, dataset, and questionnaire specifically with the UK in mind.

To ensure a high quality of results, we only allowed participants who had previously participated in five or more studies (participants had participated in around 90 studies on average) with an approval rate of 80% or higher to take part in the study (over 80% of participants had a 100% approval rate, minimum was 90%). Since all participants lived in the UK, participants were paid 2.50 GBP, which was the UK minimum wage [ukm] for 20 minutes of work (90% of the participants finished within 20 minutes).

4.4. Procedure

The study had three phases (see Figure 4): pre-stimulus, stimulus, and post-stimulus. The pre-stimulus in experiment 1 included the pre-test to assess attitudes towards immigration before the stimulus; while experiment 2 did not. The post-stimulus was identical across experiments and conditions.

Phase I: Pre-Stimulus. Participants were redirected to the Qualtrics [qua] experiment after they had selected the study in Prolific. Using meta information collected by Qualtrics, we checked that participants were using a desktop or a laptop computer, and a compatible web browser. After reading through and accepting the informed consent form, participants were informed of the study duration and their tasks. They were told that they would be asked to answer demographic questions, engage actively with a visualization, and answer a set of questions demonstrating that they understood and remembered what they had seen and learned while interacting with the visualization. We added the latter instruction to encourage participants to actively engage with the stimulus.

Participants then answered demographic questions used by models linking human values with attitudes towards immigration [DM12, DMSS14]: gender, age, highest level of education, subjective income, religion and political preferences. Participants to experiment 1 then answered the seven questions about attitudes towards immigration in a pre-test. Participants were then informed that the next screen would show the visualization and that they would have a few minutes to interact with it. They were also informed that their remaining time would be indicated in the bottom right corner of the display.

Phase II: Stimulus. Participants were randomly assigned to one of the three conditions (Exploration, Structure, and Empathy). They
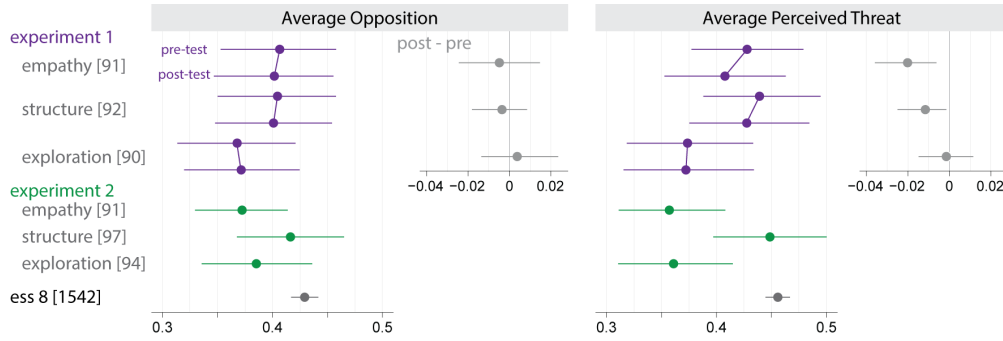
Figure 5: *Normalized sample means wiht 95% BCa CIs for the latent variables* opposition to immigration *and* perceived immigration threat *for experiment 1 (pre-test and post-test) and experiment 2 (only post-test) for the three survey conditions and the ESS run in 2016/2017.*

spent 8 minutes interacting with their assigned stimulus, as described in Section 4.1.

Phase III: Post-Stimulus. The post-stimulus phase started with three filter questions asking the color used for the migration inflow (red), the country with the greatest migration flow to the UK over the entire time (Poland), and the primary reason for migration (work). Participants were then asked to answer the 21 questions about human values and seven questions about attitudes towards immigration. This means that the immigration attitudes items are repeated-measure elements in the between-subject design for experiment 1. Each question was shown on its own page. We worded and ordered the questions like in the ESS questionnaire.

Once participants had answered all questions, they were informed that the study was designed to test the influence of visualization on attitudes towards immigration. We deliberately withheld this detail until the end to not bias the study by priming participants. Lastly, participants could leave free-form comments in a text field, then they received a code to use within Prolific to claim their financial compensation.

4.5. Hypotheses

We expected to observe differences between the three conditions, with respect to attitudes towards immigration. Specifically, our hypotheses were as follows:

$H_{empathy}$: Participants in the empathy condition will be more positive towards immigration than participants in the structure and exploration conditions. Our rationale is that the visual-narrative techniques used let participants relate to the individual fates reported.

$H_{structure}$: Participants in the structure condition will be more positive towards immigration than participants in the exploration condition. Our rationale is that a view dependent on narrated information rather than assumptions about immigration (when exploring data) might lead to more positive views, and that a structured navigation might lead one to understand the information more clearly.

4.6. Confirmatory Analysis

Before running any analysis, we discarded participants based on (a) failing to answer all filter questions, (b) answering the questionnaires too quickly, (c) selecting the same category too often, and (d) selecting often 'prefer not to say' too often. While thresholds for (b) were based on the pilot study and our own assessment, the thresholds we used for (c) and (d) were recommended by the ESS [essb]. We discarded 10% of participants of experiment 1 and 6% of experiment 2, which were distributed evenly across the conditions. While participants answered on ordinal scales, the concepts and phenomena (attitudes towards immigration and human values) behind them are considered continuous in the social science literature (e.g., [DMBS08, DMSS14]). Therefore, we treat the results as continuous variables.

For both experiments we present the results of the attitudes towards immigration variables **opposition** and **perceived threat** for each condition. For experiment 1 we also calculated the pairwise average difference between the pre-stimulus and the post-stimulus attitudes. Figure 5 shows the normalized sample means together with 95% BCa confidence intervals (CIs) based on 10,000 bootstrap replicates. For the interpretation of the statistical significance of the overlap of CI error bars we refer to [KA13].

Across both experiments, all conditions, and both variables the attitudes of participants towards immigration ranged from 0.35 to 0.45 (on a $[0-1]$ range), which means relatively positive to moderate. The range of these results is similar to UK results from the latest 2016/2017 ESS. But since the ESS is conducted in a different setting (a one-hour personal interview), we refrain from drawing conclusions from this comparison. The higher-order variables of the **human values** (e.g., openness to change and conservation) are similar across conditions and are within the range of the ESS. The detailed results of the human values and demographics are part of the supplemental material.

In experiment 1, participants in both storytelling conditions are more similar in their attitudes towards immigration than participants in the exploration condition. However, in experiment 2 participants in the empathy condition and the exploration condition are more similar than participants in the structured condition. While these tendencies can be seen in both variables, only the differences

in the perceived threat variable have some statistical significance looking at the overlap of the CI error bars.

Our confirmatory analysis does not support neither $H_{empathy}$ nor $H_{structure}$. Instead, we obtained two major results, both concerned with the *perceived threat* variable, that provide more nuanced answers to our initial questions.

Result 1: experiment 1 shows a significant but small effect of condition on perceived threat. With the pre-post-test design of experiment 1, both storytelling conditions had a small but clear effect on participants' attitude change towards immigration (their corresponding *post* − *pre* confidence intervals in Figure 5 do not cross the 0 vertical line). Participants in this two conditions perceive immigration less a threat after seeing the stimulus. While this effect can be described as *significant*, it is also small, with a point estimate smaller than 2% for both storytelling conditions.

Result 2: experiment 2 shows a large effect of condition on perceived threat. Participants in the structure condition scored higher on that scale than participants in the two other conditions. The likely difference is one order of magnitude larger than in experiment 1, with a difference of approximately 10% between the the structure condition and the other conditions.

4.7. Exploratory Analysis

To inform our discussion and provide more context to the confirmatory analysis we conducted further exploratory analysis. Specifically, we add perspective to **result 1** and **result 2** by looking at demographics and human values.

Result 1: The demographic variables shed some additional light on the observed attitude changes with regard to the average perceived threat variable in experiment 1. Following Figure 6 (top row), in the empathy design, participants that are female, in the middle age cohort, can cope or comfortably live on their income, have a center-left political orientation or are relatively more religious, showed more significant attitude change. For the structure design participants that are female, have a higher education, are relatively more religious, or have a central political orientation showed more meaningful attitude change. With regard to human values (Figure 7, top row) we observe that in the empathy design the effect is similar for both conservative and open participants. This is different for the structured condition, where only conservative participants significantly changed their attitude.

Result 2: The demographic variables provide additional context to the adversarial effect of the structure condition on the perceived threat variable in experiment 2. Following Figure 6 (bottom row), this effect applied mainly to male participants. We also observe that older participants contribute more to the observed effect than younger participants; that the effect appears to be irrespective of level of education; that those who can cope on their current income are more affected than those who find it difficult to live comfortably on their income; that more religious participants and participants with right political orientation contribute to the effect in the structured condition. With regards to human values, the effect is similar for both conservative and open participants, but mostly

self-transcendent people significantly changed their attitude. More detailed figures of the exploratory analysis, including the average opposition variable, are part of the supplemental material.

5. Discussion

In this section, we interpret these unexpected results and discuss possible explanations as well as implications for the design of visual data stories and for research in this area.

5.1. Can Anthropomorphism Elicit Empathy?

Result 1 shows a small but significant average change of immigration attitudes in the empathy condition. While this meets our expectation regarding $H_{empathy}$, the results of experiment 2 overall did not confirm this effect. This contradiction together with the small effects observed, is the second attempt after Boy et al. [BPE*17]'s study that fails at clearly demonstrating the benefits of designing for empathy in visual data storytelling. That stands in contradiction to our common and reasonable belief that empathy in visual data storytelling might affect viewers of the visualization.

It might be that using anthropomorphized data graphics does not trigger empathy as expected; but it might also be, with a narrower implication, that empathy does not influence people's attitudes towards immigration in general. The exploratory analysis in regard to result 1 showed that different demographics tend to be more influenced by the empathy design than others. On the other hand participants with differences in human values tend to be influenced equally by the empathy design. This calls for further studies of data, storytelling, anthropomorphism, empathy, and attitudes.

5.2. Could Structure Act Against Evidence-Based Understanding?

We formulated $H_{structure}$ assuming that a view dependent on evidence rather than (potentially prejudiced) assumptions about immigration might lead to more positive attitudes; and that a structured navigation through the data might lead one to understand data-evidence more clearly; which would by transitivity result in the structure condition leading to more positive attitudes towards immigration. While results from experiment 1 go in the direction of this hypothesis in terms of perceived threat, the effect size is small. With regard to Result 1, the exploratory analysis stressed that the effect depends more on demographics with the empathy design than with the structure design.

In contrast, the results for the perceived threat variable in experiment 2 indicate that participants in the structure condition perceived immigration on average as more a threat than those in the empathy and exploration conditions. Assuming homogeneity of population between the different conditions it appears that the structure condition had an effect opposite to what we expected. As the exploratory analysis showed, the effect is equally pronounced across most of the demographic variables.

The common assumption that structured navigation assists in people's understanding of facts, especially as challenged by our findings, has implications for visualization design. It is possible
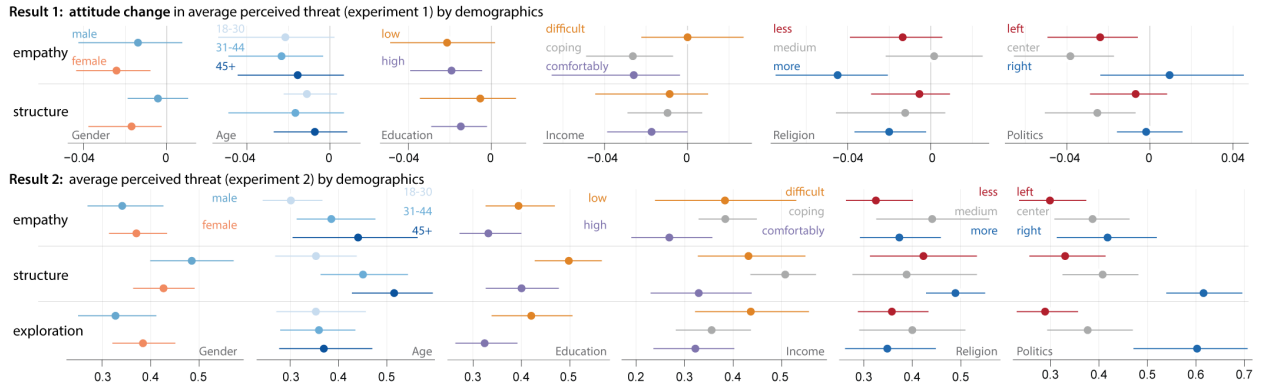
Figure 6: *Exploratory analysis for **Result 1** and **Result 2** by **demographic variable**. **Gender** compares **male** and **female** participants (none answered 'prefer not to say'). **Age** is grouped by thirds. **Education** bins Below Standards, GCSE Level Education, and A-Level Education into **low**, and Degree or Graduate Education and Post-Graduate Education into **high**. **Income** bins participants finding it very difficult or difficult to live on present income into **difficult**, and maintains the two othr original groups: **coping** and living **comfortably** on income. Due to unbalanced distribution in the binning of **Religion**, we created three bins of equal size: **less religious**, **medium religious** and **more religious**. **Politics** separates participants orientation on the **left** and the **right** spectrum, and all participants who selected the **center** option.*
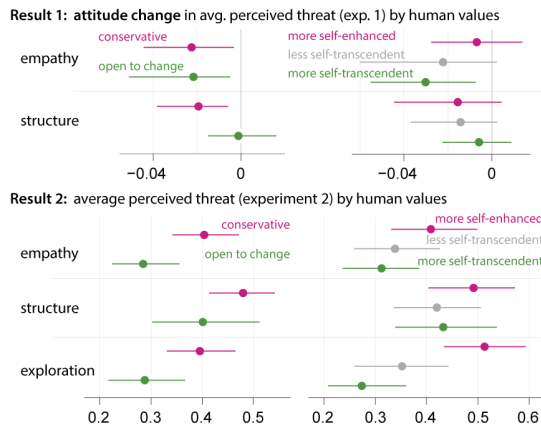

Figure 7: *Results by **higher-order human values**. One dimension contrasts **conservative** participants with people **open to change**. The second dimension, contrasting self-enhancement and self-transcendence, is a relative comparison due to unbalanced distribution. The participants are grouped in three bins: **more self-enhanced**, **less self-transcendent**, and **more self-transcendent**.*

that the author-driven approach in the structure condition might play a role in participants' responses. In contrast to the structure condition, the exploration condition lets people freely explore the data, without being constrained to the sequence decided by the author of the visualization having a narrative style that might be less perceived as patronizing. This is an avenue worth exploring in the future, especially with sensitive topics such as immigration. There is also the possibility that we introduced unconscious biases in our study, as we discuss next.

5.3. Potential Experimenter Biases

Data and (narrative) visualization can be used to communicate information in a biased way trying to influence people's attitudes. This is problematic if happening wittingly. Providing credits, provenance of the data and details of the design process can allow one to form one's own opinion. New design approaches like *Literate Visualization* [WKD18] might help to communicate design decisions effectively. While we chose not to disclose such information in our study for better control, we paid particular attention at keeping the message neutral and balanced.

In hindsight, however, it is possible that the structure design might have emphasized the migration from eastern countries. This *content bias* could explain the higher scores in the perceived threat variable in the structured condition. Another *unconscious bias* that we may have made is our assumption in our hypothesis that the evidence provided in the visual-narrative designs suggests that migration is not a threat. This bias is associated with the content bias, as we have made these assumptions on the basis of a balanced and neutral narrative of both visual-narrative designs and what we expected the techniques used could afford.

Finally, for empirical testing visual-narrative stimuli have to be complex enough to imitate realistic settings, which could lead to biased complexities across stimuli influencing understanding. However, the results of the filter questions (used as a proxy) do not indicate that the different designs had an impact on understanding.

5.4. Measuring the Effects of Storytelling is Challenging

Empirical testing of the effects of visual-narrative techniques is a balancing act. Testing too isolated aspects out of context might reduce the relevance of any results, which makes long, complex, and realistic enough stimuli important. For a better control, especially when assessing attitudes rather than perceptual measures or preferences, a between-subject design where participants interact with

just one stimulus is beneficial. It would be hard to isolate and quantify any learning effects induced by a within-subject design.

Nevertheless, based on our results, our study joins the growing set of studies investigating the role of visual data storytelling that tend to find small or unexpected effects when testing different visual-narrative techniques or visual communication modes. However, the size of differences should be put in their context. Attitudes towards immigration are believed to be strong attitudes (research in attitude strength distinguishes between strong and weak attitudes [MH15]). Consequently, it is unlikely to observe large effects on these strongly held attitudes and small but notable effects, in this context, might be more important than they look.

In a wider context, the authoring, and the perception, of visual narratives are subjective [TPWC17]. The same data, observations, and facts often allow different interpretations and yield different narratives [CM11, Phi12] in the form of different designs or in the form of different mental images. This is fully acceptable and can be seen as a strength of narrative approaches and as an advantage to visualization designers; but it also makes it difficult to quantify effects in experimental setups, and even more to generalize results.

5.5. The Need to Venture into Other Methods

The different results of the post-stimulus assessments between the experiments can have various reasons. It could simply be (1) statistical noise, (2) that the two sampled groups of participants vary to a large extent (e.g., in their prior beliefs), or (3) due to the different experimental design, i.e., people respond differently because they are asked the question twice - or a mix of those reasons. Assuming the pre-test can influence participants, the experimental design 1 is in a way inferior to the experimental design 2 where participants were not primed. It is superior in other ways (e.g., more statistical power). The reasons why participants of experiment 1 were not prompted to change their minds could be that: (a) they were not influenced by the stimuli; or (b) that they were subject to conservatism bias. Our second experiment, where we only assessed post-stimulus attitudes, resulted in different attitudes towards immigration compared to the post-test in the first experiment, suggesting that including a pre-test indeed *does* affect the results.

In the confirmatory analysis we did not adjust for demographic or other variables (e.g., human values) due to the large number of participants and the random group allocation, as recommended in health sciences [Sen13, MPP18] for example. Nevertheless, we cannot exclude the possibility that differences observed in experiment 2 are due to differences in prior beliefs and personal views rather than the visual-narrative techniques, which would be indicated by experiment 1. The explanatory results suggest that there are small differences in the demographic variables (Figure 6) thus further research should investigate these variables in isolation.

We see our results as a warning bell for future studies in this area. Since it is difficult in practice to isolate the impact of visual narratives from the context around them, we, as a community, need to consider developing alternative approaches when assessing attitudes and beliefs in narrative visualization to the widely employed designs borrowed from highly controlled perceptual studies. For example, recent work [NBS18] struck a new path by conducting

micro-phenomenological interviews, to target experience in visual data storytelling. The application of such or similar qualitative evaluation approaches (e.g., case studies, focus group discussions, the use of additional media and narratives on the same topic) to the visual narratives presented, which include techniques such as anthropomorphic elements or structured sequential revelations, can be a next valuable research contribution.

5.6. Limitations and Future Work

As any study attempting to quantify some human characteristics, our work has several limitations. In both visual-narrative designs, we used text-based narration, which competes with other visual elements. Audio-based narration would reduce the load on the visual channel, which could lead to other results. But the online crowdsourcing approach does not allow to control over an individual's setup (e.g, over the noise level, or audio quality). Future 'in-house' experiments could make use of the audio channel to investigate effects in multimedia data storytelling.

In a world without monetary constraints, we could have implemented additional conditions. We could have created other designs using narrative techniques targeting the same categories as we used (structure and empathy) to confirm our findings. We also could have included designs implementing narrative techniques that target other categories [RHDC18] (e.g., framing, argument, engagement/immersion) to test their influences on attitudes. Further, alternative study designs might lead to stronger conclusions. For example, combining the empathy and the structure designs as an additional condition might bring an improved significance to our results. A design that compare user interfaces that either inform or persuade participants with the assessment of attitudes and prior knowledge could also add value to our findings.

Result 2 – the adversarial effect we observed in the structured condition in experiment 2 – points toward additional research on different narrative structuring techniques and their influence on attitudes. Such studies could for example facilitate work on narrative sequencing preferences [HDR*13, HKL17] or argument structure [Kos17]. Additional studies are required before we can generalize our findings and draw broader conclusions and guidelines for designing visual narratives that can influence people's attitudes.

6. Conclusions

Compared to the popularity and the widespread use of visual data storytelling, visualization research analyzing the mechanisms of storytelling is underrepresented. With the first study within our field testing if different visual-narrative techniques can influence attitude, we contribute to improving our overall understanding of visual data storytelling. We found no evidence to support our initial assumptions that empathy-evoking and structured narrative visualizations could elicit more positive attitudes towards immigration than non-narrative displays. What we as a visualization community expect storytelling to afford is yet to be demonstrated empirically as the storytelling modes we investigated did not strongly influence people's attitudes. Our results also have methodological implications. When assessing attitudes that tend to be strongly held,

priming participants through pre-stimulus assessments is likely to be problematic as it might introduce conservatism bias.

(Negative) results such as ours, which do not confirm our hypotheses nor support common assumptions, are important contributions to our research field as they prevent us from relying on plausible but unchallenged assumptions. The field of research concerned with visual data storytelling in particular, and visualization in general, is still in its early stages and many more studies need to investigate our often unchallenged assertions in this field. We stress the need to gather more empirical evidence before making strong claims regarding the benefits of storytelling in visualization – as it appears that benefits that might have been demonstrated in other disciplines do not necessarily apply to the field of visualization.

These findings help to better anticipate the effects of narrative techniques used in visualization over people's attitudes towards topics regarded as contentious and strongly held. We contribute new building blocks to the growing knowledge of storytelling in visualization, which can be used to design further experiments investigating the use of visual-narrative techniques.

References

[AHRL*15] AMINI F., HENRY RICHE N., LEE B., HURTER C., IRANI P.: Understanding Data Videos: Looking at Narrative Visualization through the Cinematography Lens. In *Proc. CHI '15* (2015), pp. 1459–1468. doi:10.1145/2702123.2702431. 2

[ARL*17] AMINI F., RICHE N. H., LEE B., MONROY-HERNANDEZ A., IRANI P.: Authoring Data-Driven Videos with DataClips. *IEEE TVCG 23*, 1 (2017), 501–510. doi:10.1109/TVCG.2016.2598647. 2

[ARL*18] AMINI F., RICHE N. H., LEE B., LEBOE-MCGOWAN J., IRANI P.: Hooked on Data Videos: Assessing the Effect of Animation and Pictographs on Viewer Engagement. In *Proc. AVI '18* (2018), pp. 21–29. doi:10.1145/3206505.3206552. 2

[BBK*16] BORKIN M. A., BYLINSKII Z., KIM N. W., BAINBRIDGE C. M., YEH C. S., BORKIN D., PFISTER H., OLIVA A.: Beyond Memorability: Visualization Recognition and Recall. *IEEE TVCG 22*, 1 (2016), 519–528. doi:10.1109/TVCG.2015.2467732. 1, 2, 4

[BDF15] BOY J., DETIENNE F., FEKETE J.-D.: Storytelling in Information Visualizations: Does it Engage Users to Explore Data? In *Proc. CHI '15* (2015), pp. 1449–1458. doi:10.1145/2702123.2702452. 1, 2, 3, 4

[BLB*17] BREHMER M., LEE B., BACH B., RICHE N. H., MUNZNER T.: Timelines Revisited: A Design Space and Considerations for Expressive Storytelling. *IEEE TVCG 23*, 9 (2017), 2151–2164. doi:10.1109/TVCG.2016.2614803. 2

[BLV*17] BREHMER M., LEE K., VIOLA I., SEO J., LEE B.: Demonstrating the Value of Visualization: Highlights from the 2017 PacificVis Visual Data Storytelling Contest. In *Poster Proceedings of the 2017 IEEE VIS Conference* (2017). 1

[BPE*17] BOY J., PANDEY A. V., EMERSON J., SATTERTHWAITE M., NOV O., BERTINI E.: Showing People Behind Data: Does Anthropomorphizing Visualizations Elicit More Empathy for Human Rights Data? In *Proc. CHI '17* (2017), pp. 5462–5474. doi:10.1145/3025453.3025512. 1, 2, 3, 8

[BRCP17] BACH B., RICHE N. H., CARPENDALE S., PFISTER H.: The Emerging Genre of Data Comics. *IEEE CG&A 38*, 3 (2017), 6–13. 2

[BVB*13] BORKIN M., VO A., BYLINSKII Z., ISOLA P., SUNKAVALLI S., OLIVA A., PFISTER H.: What Makes a Visualization Memorable? *IEEE TVCG 19*, 12 (2013), 2306–2315. doi:10.1109/TVCG.2013.234. 1, 2, 4

[BW14] BADAWOOD D., WOOD J.: The Effect of Information Visualization Delivery on Narrative Construction and Development. In *Proc. EuroVis '14* (2014). 2, 3

[chi] CHI. https://chi2019.acm.org/. 1

[CM11] CRUZ P., MACHADO P.: Generative Storytelling for Information Visualization. *IEEE CG&A 31*, 2 (2011), 80–85. 2, 10

[CRS*20] CONCANNON S., RAJAN N., SHAH P., SMITH D., URSU M., HOOK J.: Brooke leave home: Designing a personalized film to support public engagement with open data. In *Proceedings of the ACM CHI 2020 Conference on Human Factors in Computing Systems* (1 2020), Association for Computing Machinery (ACM). Authors. This is an author-produced version of the published paper. Uploaded in accordance with the publisher's self-archiving policy. Further copying may not be permitted; contact the publisher for details. 2, 3

[CSF*19] CHOE E. K., SAKAMOTO Y., FATMI Y., LEE B., HURTER C., HAGHSHENAS A., IRANI P.: Persuasive data videos: Investigating persuasive self-tracking feedback with augmented data videos. In *Proceedings of AMIA 2019* (November 2019). 2

[DBD17] DIMARA E., BEZERIANOS A., DRAGICEVIC P.: Narratives in Crowdsourced Evaluation of Visualizations: A Double-Edged Sword? In *Proc. CHI '17* (New York, NY, USA, 2017), ACM, pp. 5475–5484. doi:10.1145/3025453.3025870. 2, 3

[Den16] DENIL M.: Storied Maps. *Cartographic Perspectives*, 84 (2016), 5–22. doi:10.14714/CP84.1374. 1

[Dia10] DIAKOPOULOS N.: Game-y Information Graphics. In *Proc. CHI EA '10* (2010), pp. 3595–3600. doi:10.1145/1753846.1754024. 1, 2, 3, 5

[DKN11] DIAKOPOULOS N., KIVRAN-SWAINE F., NAAMAN M.: Playable Data: Characterizing the Design Space of Game-y Infographics. In *Proc. CHI '11* (2011), pp. 1717–1726. doi:10.1145/1978942.1979193. 2

[DM12] DAVIDOV E., MEULEMAN B.: Explaining Attitudes Towards Immigration Policies in European Countries: The Role of Human Values. *Journal of Ethnic and Migration Studies 38*, 5 (2012), 757–775. doi:10.1080/1369183X.2012.667985. 3, 6

[DMBS08] DAVIDOV E., MEULEMAN B., BILLIET J., SCHMIDT P.: Values and Support for Immigration: A Cross-Country Comparison. *European Sociological Review 24*, 5 (2008), 583–599. doi:10.1093/esr/jcn020. 3, 7

[DMSS14] DAVIDOV E., MEULEMANN B., SCHWARTZ S. H., SCHMIDT P.: Individual Values, Cultural Embeddedness, and Anti-Immigration Sentiments: Explaining Differences in the Effect of Values on Attitudes Toward Immigration Across Europe. *Kölner Zeitschrift für Soziologie und Sozialpsychologie 66*, 1 Supplement (2014), 263–285. doi:10.1007/s11577-014-0274-5. 3, 6, 7

[DS12] DILLARD J. P., SHEN L.: *The SAGE Handbook of Persuasion: Developments in Theory and Practice*, second ed. SAGE Publications, Thousand Oaks, California, 2012. 3

[Dua10] DUARTE N.: *Resonate: Present Visual Stories That Transform Audiences*. Wiley, Hoboken, New Jersey, 2010. 1

[essa] Antecedents of Anti-Immigration Attitudes. http://essedunet.nsd.uib.no/cms/topics/immigration/1/2.html. 3

[essb] European Social Survey (ESS). http://www.europeansocialsurvey.org/. 1, 3, 6, 7

[Few17] FEW S.: Tell Me a Story, or Not. https://www.perceptualedge.com/blog/?p=2568, 2017. 1

[Fig14a] FIGUEIRAS A.: How to Tell Stories Using Visualization. In *Proc. IV '14* (2014), pp. 18–26. doi:10.1109/IV.2014.78. 1, 2

[Fig14b] FIGUEIRAS A.: Narrative Visualization: A Case Study of How to Incorporate Narrative Elements in Existing Visualizations. In *Proc. IV '14* (2014), pp. 46–52. doi:10.1109/IV.2014.79. 2

[GP01] GERSHON N., PAGE W.: What Storytelling can do for Information Visualization. *Communications of the ACM 44*, 8 (2001), 31–37. 2

[HD11] HULLMAN J., DIAKOPOULOS N.: Visualization Rhetoric: Framing Effects in Narrative Visualization. *IEEE TVCG 17*, 12 (2011), 2231–2240. 2

[HDR*13] HULLMAN J., DRUCKER S., RICHE N. H., LEE B., FISHER D., ADAR E.: A Deeper Understanding of Sequence in Narrative Visualization. *IEEE TVCG 19*, 12 (2013), 2406–2415. 1, 2, 3, 5, 10

[HJR10] HERMAN D., JAHN M., RYAN M. (Eds.): *Routledge Encyclopedia of Narrative Theory*. Routledge, London, 2010. 2

[HKL17] HULLMAN J., KOSARA R., LAM H.: Finding a Clear Path: Structuring Strategies for Visualization Sequences. In *Proc. EuroVis '17* (2017), Heer J., Ropinski T., van Wijk J., (Eds.). 2, 10

[HRR20] HEYER J., RAVEENDRANATH N. K., REDA K.: Pushing the (visual) narrative: the effects of prior knowledge elicitation in provocative topics. URL: http://par.nsf.gov/biblio/10133071, doi:10.1145/3313831.3376887. 2

[iib] Information is Beautiful Awards. https://www.informationisbeautifulawards.com/. 1

[JRFE07] JOWELL R., ROBERTS C., FITZGERALD R., EVA G.: *Measuring Attitudes Cross-Nationally*. SAGE Publications, London, 2007. doi:10.4135/9781849209458. 1, 3

[JSM*16] JENNY B., STEPHEN D. M., MUEHLENHAUS I., MARSTON B. E., SHARMA R., ZHANG E., JENNY H.: Design Principles for Origin-Destination Flow Maps. *Cartography and Geographic Information Science 45*, 1 (2016), 62–75. doi:10.1080/15230406.2016.1262280. 4

[KA13] KRZYWINSKI M., ALTMAN N.: Points of significance: Error bars. *Nature Methods 10*, 10 (10 2013), 921–2. Copyright - Copyright Nature Publishing Group Oct 2013; Last updated - 2014-04-10. URL: https://0-search-proquest-com.wam.city.ac.uk/docview/1458601290?accountid=14510. 7

[KLK18] KONG H.-K., LIU Z., KARAHALIOS K.: Frames and Slants in Titles of Visualizations on Controversial Topics. In *Proc. CHI '18* (2018), pp. 438–438. doi:10.1145/3173574.3174012. 2

[Kos17] KOSARA R.: An Argument Structure for Data Stories. In *Proc. EuroVis '17 Short Papers* (2017). 2, 10

[LRIC15] LEE B., RICHE N. H., ISENBERG P., CARPENDALE S.: More Than Telling a Story: Transforming Data into Visually Shared Stories. *IEEE CG&A 35*, 5 (2015), 84–90. 1, 2

[MF17] MOCNIK F.-B., FAIRBAIRN D.: Maps Telling Stories? *The Cartographic Journal 55*, 1 (2017), 36–57. doi:10.1080/00087041.2017.1304498. 4

[MH15] MAIO G. R., HADDOCK G.: *The Psychology of Attitudes & Attitude Change*, second ed. SAGE Publications, Los Angeles, 2015. 3, 10

[ML13] MARKAKI Y., LONGHI S.: What Determines Attitudes to Immigration in European Countries? An Analysis at the Regional Level. *Migration Studies 1*, 3 (2013), 311–337. doi:10.1093/migration/mnt015. 3

[MLF*12] MA K.-L., LIAO I., FRAZIER J., HAUSER H., KOSTIS H.-N.: Scientific Storytelling Using Visualization. *IEEE CG&A 32*, 1 (2012), 12–19. 1, 2

[MPP18] MUTZ D. C., PEMANTLE R., PHAM P.: The perils of balance testing in experimental design: Messy analyses of clean data. *The American Statistician* (2018), 1–11. doi:10.1080/00031305.2017.1322143. 10

[MRL*17] MCKENNA S., RICHE N. H., LEE B., BOY J., MEYER M.: Visual Narrative Flow: Exploring Factors Shaping Data Visualization Story Reading Experiences. *Computer Graphics Forum 36*, 3 (2017), 377–387. 1, 2, 5

[Mue12] MUEHLENHAUS I.: If Looks Could Kill: The Impact of Different Rhetorical Styles on Persuasive Geocommunication. *Cartographic Journal 49*, 4 (2012), 361–375. doi:10.1179/1743277412Y.0000000032. 1, 2, 3

[nap] NAPA Cards: Narrative Patterns for Data Stories. http://napa-cards.net/. 2, 4, 5

[NBS18] NOWAK S., BARTRAM L., SCHIPHORST T.: A micro-phenomenological lens for evaluating narrative visualization. In *2018 IEEE Evaluation and Beyond - Methodological Approaches for Visualization (BELIV)* (Oct 2018), pp. 11–18. doi:10.1109/BELIV.2018.8634072. 10

[Nis94] NISSANI M.: Conceptual conservatism: An understated variable in human affairs? *The Social Science Journal 31*, 3 (1994), 307–318. doi:https://doi.org/10.1016/0362-3319(94)90026-4. 5

[NK15] NUSSBAUMER KNAFLIC C.: *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Wiley, Hoboken, New Jersey, 2015. 1

[nyt] The New York Times. https://www.nytimes.com. 1

[nzz] NZZ Storytelling. https://www.nzz.ch/storytelling/. 1

[O'K16] O'KEEFE D. J.: *Persuasion: Theory and Research*, third ed. SAGE Publications, Thousand Oaks, California, 2016. 2, 3

[ons] ONS - Office for National Statistics. https://www.ons.gov.uk/. 4

[ORS09] OECHSSLER J., ROIDER A., SCHMITZ P. W.: Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization 72*, 1 (2009), 147–152. doi:https://doi.org/10.1016/j.jebo.2009.04.018. 5

[ovc] OpenVisConf. http://openvisconf.com. 1

[Phi12] PHILLIPS J.: Storytelling in Earth Sciences: The Eight Basic Plots. *Earth-Science Reviews 115*, 3 (2012), 153–162. doi:10.1016/j.earscirev.2012.09.005. 10

[PMN*14] PANDEY A. V., MANIVANNAN A., NOV O., SATTERTHWAITE M., BERTINI E.: The Persuasive Power of Data Visualization. *IEEE TVCG 20*, 12 (2014), 2211–2220. doi:10.1109/TVCG.2014.2346419. 1, 2, 3

[pro] Prolific. https://prolific.ac/. 6

[qua] Qualtrics. https://www.qualtrics.com/. 6

[RHDC18] RICHE N. H., HURTER C., DIAKOPOULOS N., CARPENDALE S.: *Data-Driven Storytelling*. CRC Press, Boca Raton, 2018. 1, 4, 5, 10

[Rot] ROTH R.: Cartographic Design as Visual Storytelling. http://bit.ly/rothstorytelling. 2

[Rya04] RYAN M.-L.: *Narrative Across Media: The Languages of Storytelling*. Frontiers of narrative. University of Nebraska Press, Lincoln, 2004. 2

[Sch07] SCHWARTZ S. H.: Value Orientations: Measurement, Antecedents and Consequences Across Nations. In *Measuring Attitudes Cross-Nationally*. SAGE Publications, London, 2007, pp. 169–203. doi:10.4135/9781849209458.n9. 3

[Sch17] SCHWABISH J.: More Story References and Resources. https://policyviz.com/2017/03/24/more-story-references-and-resources/, 2017. 1

[Sen13] SENN S.: Seven Myths of Randomisation in Clinical Trials. *Statistics in Medicine 32*, 9 (2013), 1439–1450. doi:10.1002/sim.5713. 10

[SH10] SEGEL E., HEER J.: Narrative Visualization: Telling Stories with Data. *IEEE TVCG 16*, 6 (2010), 1139–1148. doi:10.1109/TVCG.2010.179. 1, 2

[SLRS16] STOLPER C. D., LEE B., RICHE N., STASKO J.: *Emerging and Recurring Data-Driven Storytelling Techniques: Analysis of a Curated Collection of Recent Stories*. Tech. Rep. MSR-TR-2016-14, Microsoft Research, 2016. 2

[tap]   Tapestry Conference. http://www.tapestryconference.com/. 1

[tim]   Timeline Storyteller. https://timelinestoryteller.com/. 2

[TPWC17]   THUDT A. F., PERIN C., WILLETT W. C., CARPENDALE
    S.: Subjectivity in Personal Storytelling with Visualization. *Information Design Journal 23*, 1 (2017), 48–64. doi:10.175/idj.23.1.
    07thu. 10

[ukm]   National Minimum Wage and National Living Wage rates -
    GOV.UK. https://www.gov.uk/national-minimum-wage-rates. 6

[vis]   IEEE VIS. http://ieeevis.org/. 1

[WKD18]   WOOD J., KACHKAEV A., DYKES J.: Design exposition
    with literate visualization. *IEEE TVCG 25*, 1 (2018). doi:10.1109/
    TVCG.2018.2864836. 9