Showing People Behind Data: Does Anthropomorphizing Visualizations Elicit More Empathy for Human Rights Data?

Jeremy Boy
UN Global Pulse, USA
myjyby@gmail.com

Anshul Vikram Pandey
NYU Tandon, USA
anshul.pandey@nyu.edu

John Emerson
NYU School of Law, USA
john@backspace.com

Margaret Satterthwaite
NYU School of Law, USA
satterth@exchange.law.nyu.edu

Oded Nov
NYU Tandon, USA
onov@nyu.edu

Enrico Bertini
NYU Tandon, USA
enrico.bertini@nyu.edu

ABSTRACT
We investigate the impact of using anthropomorphized data graphics over standard charts on viewers' empathy for, and prosocial behavior toward suffering populations, in the context of human rights narratives. We present a series of experiments conducted on Amazon Mechanical Turk, in which we compare various forms of anthropomorphized data graphics—ranging from a single human figure that 'fills up' to show proportional data, to separated groups of individual human beings—with a standard chart baseline. Each experiment uses two carefully crafted human rights data-driven stories to present the graphics. Contrary to our expectations, we consistently find that anthropomorphized data graphics and standard charts have very similar effects on empathy and prosocial behavior.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords
Information Visualization for the People; Anthropographics; Empathy; Prosocial Behavior; Human Rights

INTRODUCTION
In this article, we investigate the growing assumption that visually connecting abstract data with iconic representations of people can elicit empathy for, and encourage prosocial behavior toward those people [36, 42, 55, 66]. We focus specifically on data visualizations of human rights/humanitarian (*HR*) issues in data-driven stories, as these generally describe people's plight, and have real potential for eliciting empathy. The work we present is the result of a collaboration between Information Visualization (Infovis) and HR researchers.

CHI 2017, May 06–11, 2017, Denver, CO, USA.

DOI: http://dx.doi.org/10.1145/3025453.3025512

Many HR practitioners are concerned that describing large scale atrocities using abstract data might create "compassion fatigue" [28, 60] by distancing readers from the reality of human suffering. HR-related data are often very sensitive regarding *e.g.,* the privacy, security, or safety of entire populations, and can seldom be presented publicly without some amount of aggregation. This makes it impossible for HR advocates to use re-identifiable information like portrait photos to humanize the data—even thought photos are known to elicit affective responses [30]—and forces them to seek out other strategies for illustrating the "human dimension" [28] of abstract data.

Visualization designers often use *anthropomorphized* data graphics [66] (or *anthropographics*, see [1]) to this end. The common rationale is that because such visualizations create an immediate visual connection between abstract data and actual people [36], they are better suited than standard charts for eliciting empathy and prosocial behavior [42, 66]. We refer to this rationale as the *anthropographic assumption* [36]. Although some critics are skeptical of the need for empathy in visualization (*e.g.,* [17]), others believe it can be particularly useful for leveraging public awareness in the pursuit of social change [31, 42, 66]. Many have discussed and nourished the anthropographic assumption [17, 36, 55, 58, 67], but there is no empirical evidence to support it.

Here, we formalize a design space for anthropographics, and contribute the results of seven experiments as an initial assessment of the anthropographic assumption in the context of HR narratives. In contrast to our expectations, we find that anthropographics and standard charts (*e.g.,* pie charts) have very similar effects on empathy and prosocial behavior. We conclude that for HR narratives, anthropographics are neither truly beneficial, nor detrimental. This complements Bateman *et al.*'s call to learn more about the effects of different types of visual embellishment in charts [6], and opens new perspectives for exploring the benefits of anthropographics.

BACKGROUND
In this section, we first develop the arguments in favor of the anthropographic assumption. We then motivate the general design of our experiments with prior work on empathy, and introduce known mechanisms behind a common form of prosocial behavior: charitable giving.

Anchoring Graphics in Empathy

In January 2015, Jake Harris posted an article discussing the importance of establishing a connection between data and actual people in data-driven stories [36]. He argues that "from a distance," *i.e.,* when abstract representations are used to present human-related data, "it's easy to forget the dots are people." He claims that "connecting with the dots" can elicit empathy, which is "important for any dataset [used] to report data about people or that affects people."

Lambert, Rees, and Zer-Aviv all agree with this assumption [42, 55, 66]. While Harris mostly suggests using "wee people" to "anchor graphics in empathy" [36], Rees encourages visualization designers to think about the "atom" of the data, *e.g.,* an individual person; to "get into the life" of that person; to "speak to it"; and to let the "greater piece" grow out of it [55]. Evidence in other fields (*e.g.,* robotics and animal rights) provides support for these approaches, as it indicates the more something appears humanlike, the more people feel empathy for it [37, 54]. From a semiotics perspective, using "wee people"—instead of only textual annotations on top of standard charts—presumably has the advantage of eliciting visual interpretation mechanisms, which are far more open than those of verbal texts [49]. Visual information typically builds on experiential knowledge [52], whereas text relies more on cultural conventions [32]. More *realistic* and/or *expressive* visual renderings of victims [45] may tap into this experiential knowledge, and affect viewers at a preverbal level [49].

Lester *et al.* have found that adding lifelike characters to interactive learning environments has a strong positive effect on students' affect, motivation, and perception of their experience [44]. They call this the *persona effect*, and argue that *expressive* characters are most effective. However, Miksatko *et al.* have failed to reproduce this effect, and conclude that such characters are neither detrimental, nor truly beneficial [48]. Genevsky *et al.* have found that showing pictures of orphan children to potential charity-donors has a positive effect on their affect, which in turn drives them to give more money [30]. They link this to the *identifiable victim effect*, which broadly describes the belief that people empathize more, and are willing to expend greater resources to help identified victims than to help equal, or greater numbers of *statistical victims* [39]. While their results show that photo portraits are more effective than abstract silhouettes, the iconic human representations they use are not very *expressive*, and one could argue that using "wee people" instead of standard charts might better help reproduce the *identifiable victim effect*, and thus encourage more prosocial behavior [42, 66]. Establishing this is important, as photos are not always available, and are often difficult to use in the context of HR data-driven advocacy—whether simply because the data are aggregated beyond the point of individual distinction, or because they are too sensitive.

Finally, despite the "chart junk" debate [6, 62], previous work in Infovis and HCI has shown that the use of pictorial information like "wee people" in visualizations has a number of benefits. Borgo *et al.* have found that visual embellishments help viewers grasp key concepts more effectively [12], *e.g.,* the connection between abstract data and actual people; Borkin *et al.* argue that pictograms can improve recognition and memorability [14], and that creating redundancy, *e.g.,* between text and graphics, can help convey messages more effectively [13]; and Haroz *et al.* have shown that Isotype-like visualizations [2, 16] are generally engaging and can facilitate recall [35].

Empathy

Although Cairo, Harris, Lambert, Rees, Schwabish, and Zer-Aviv all speak of eliciting empathy with visualizations [17, 36, 42, 55, 66], it is sometimes unclear whether they share the same definition of *empathy*. The term often suffers from inconsistent use [27], and is commonly interchanged with *e.g., sympathy*. This creates confusion [11], which we attempt to clarify here for the purpose of our experiments.

Empathy is broadly defined as "an affective response appropriate to someone else's situation rather than one's own" [7]. It is generally considered to have two dimensions: an affective dimension, and a cognitive dimension [7, 10, 21, 27, 38, 47, 64]. *Affective empathy* relates to the vicarious experiencing of another's emotional state. *Cognitive empathy* relates to one's ability to accurately imagine another's internal states (*e.g.,* her thoughts, feelings, or intentions) [38].

Affective empathy has two components which lead to *empathic distress* [9, 21, 25, 27, 38]: empathic concern and personal distress. *Empathic concern* describes *other-oriented* feelings of compassion and sympathy [64], which are not necessarily congruent with the feelings of the distressed other. It relates to feeling *for* someone, rather than to feeling *with* someone. *Personal distress* on the other hand describes *self-oriented* feelings of anxiety, discomfort, and unease triggered in response to the distress of others. In this article, any generic use of the term *empathy* refers to these two emotional components.

Hoffman identifies five "modes" of empathy-arousal [38]. *Mimicry*, *classical conditioning*, and *direct association* are primitive, automatic, and preverbal modes. Presumably, *realistic* and/or *expressive* visual renderings of victims [45] should help trigger these. *Verbally mediated association* and *perspective-taking* are higher-level cognitive modes that give scope to one's feelings, and allow to empathize with others who are not present. They are central to *narrative empathy*, *i.e.,* "the sharing of feeling and perspective-taking induced by reading, viewing, hearing, or imagining narratives of another's situation or condition" [40], and suggest that empathy can be aroused simply by conceiving of another's plight [38], and that it can be generalized to a type of plight, or to the plight of an entire population (*e.g.,* the poor or oppressed) [7].

Finally, empathy is often considered a precursor and motivator for prosocial behavior [47, 64]. Batson's *empathy-altruism hypothesis* [7] typically suggests that observers will seek to relieve their empathic distress by helping others. However, this is not always true. Observers may enjoy another's misfortune if their prior relationship is bad [10]. Egoistic motives may interfere [8], and the distress may become so aversive that observers end up focusing solely on temselves [38]. Empathy also has a *familiarity bias* [38]: observers generally tend to have higher feelings for kin, in-group members, family, and people who share their personal concerns.

Measuring Empathy and Understanding Donating

Research in developmental and social psychology has studied empathy either as a mental state that can be triggered [8, 24], or as a personal disposition that develops during childhood [24, 47]. Here, we mainly focus on the prior type of studies.

Batson uses a procedure which consists in prompting participants with a fictional scenario describing a person in need, and asking them to report on a 7-point scale how strongly they feel different emotions described in a list of *emotion adjectives* [7]. This list includes eight adjectives assumed to reflect personal distress, and six assumed to reflect empathic concern. Similarly, the *picture/story* procedure [27] consists in telling participants brief stories while showing them pictures (*e.g.,* photos or drawings) of protagonists in emotional situations, and asking them to verbally indicate how they feel. Our experiments are inspired by these procedures.

Meanwhile, Genevsky *et al.* have shown how empathy can shift donation preferences [30]. Lee *et al.* and Verhaert & Van den Poel have also found that empathic concern positively influences donations [43, 64]. However, Einolf argues that this may not be true when the recipients of the donation are not immediately present [24]. There are a number of factors that influence donating behavior, which may complexify the relation between empathy and donation decisions. Here, we refer to these factors as *donation biases* for convenience.

Donating is usually seen as a prosocial behavior [43, 57], for which moral identity is a strong predictor. *Moral identity* relates to the importance one attributes to traits like fairness, justice, or kindness [43]. Although it is usually theorized to increase charitable giving, Lee *et al.* have found that it can have a negative effect if the recipients of the charity are perceived as responsible for their plight [43]. They also found that *empathy* and the perception of *justice* can have mediating effects. In our experiments, we consider these factors as possible *moral*, *responsibility*, and *justice biases*. Another phenomenon that influences moral judgement—and thereby donating behavior—is proportion dominance [5, 59]. *Proportion dominance* describes the fact that people are likely to choose to help higher proportions of statistical victims than absolute numbers of them. Jenni & Lowenstein claim that this dominance is responsible for the identifiable victim effect [39]. In our experiments, we consider the possibility of a *proportion dominance bias*. Finally, *availability* has a strong influence on decision-making [63], as it operates on the notion that if something can be recalled, it must be important. We pay particular attention to *topic availability* in our experiments, as we expect it may influence the *familiarity bias*, as well as people's perception of how important it is to donate to specific causes.

GENERAL MATERIALS AND METHODS

In this section, we present the general design of our experiments. We detail the HR narratives we created, and describe our design space for anthropographics.

General Study Design

Each experiment consisted of two HR narratives. These were composed of a single chart condition, set in a unique data-driven story. Chart conditions were: an *anthropographic* (variable), and a *standard chart* (baseline)—except in Experiment 4 (*Exp.* 4), where we compared the *standard chart* with *text alone*. We set these conditions in data-driven stories for ecologic validity, as most data graphics end up embedded within additional content like text [35]—especially in HR narratives. Participants were thus exposed to one chart condition in a first story, then to the other in a different story. We opted for this design, rather than for a full-factorial design—in which participants would have seen both conditions twice (once in each story)—because we did not want to force them to read the same story several times. We feared this might create lassitude or frustration, which could have confounded their feelings. Finally, we counterbalanced the order of conditions and stories to minimize the possible effects of *anthropomorphic carry-over*. As both stories were about human suffering, we feared that participants who were exposed to the anthropographic condition in the first narrative might automatically transfer the "human dimension" [28] of the data to the second narrative, thus confounding the effect of chart conditions.

Stories

Although we only used two stories in each experiment, we initially created four. All were related to the Syrian crisis, and were developed with two HR researchers (co-authors of this article). We focused on Syria, as we expected participants would be familiar with the context. Each story described children's plight (similarly to [30]) through the lens of one of the following topics: access to clean **water**, access to **education**, **poverty**, and **internal displacement** (IDP). We felt that focusing on children would mitigate the *responsibility* and *justice biases*, since children presumably cannot be held accountable for their situation. We also expected this would amplify the *moral bias*, since helping children should be perceived as at least as morally important as helping adults, if not more. All stories shared the same five-step narrative structure, to avoid possible *narrative-structure-related* confounds. Step 1 introduced the context. Step 2 showed a geographic overview of the situation. Step 3 presented a simple demographic statistic before the crisis. Step 4 showed the same statistic today. Step 5 concluded with a takeaway consideration on how children are suffering from the situation.

Visualization template: we created a custom *slideshow* visualization template [56] to comply with our five-step narrative structure. Each *slide* corresponded to one step. Progression through the different slides was achieved using the arrow keys on a keyboard. We kept visual compositions quasi-identical for each story. Slide 1 showed only text. Slide 2 showed a choropleth map of Syria and a short paragraph of text. Slide 3 showed a chart and a short paragraph of text. Slide 4 showed the updated chart and a short paragraph of text. Slide 5 showed only text. Varying chart conditions were restricted to slides 3 and 4. We also kept other visual design attributes like fonts, background color, and transition types consistent across all stories. The only slight variations were *dominant color*—used to visually differentiate topics—and *placement* of the short paragraphs in slides 3 and 4—depending on the chart condition. We hoped this would prevent any kind of *aesthetic-preference-related* confounds. To finish, we added a simple hover interaction to the charts for showing labels and values.

Data: for our chart conditions (slides 3 and 4), we used demographic data at two different time-steps. These were percentages of the Syrian population affected by the story's topic, before the crisis began and today. We used such simple statistics for three reasons. First, because we wanted aggregated human-related data, *i.e.,* non-identifiable data, that could easily be visualized as groups of individuals (*e.g.,* 100% = 100 individuals). Second, because we believed using only proportional data—instead of using proportions and absolute numbers—would help avoid possible confounds linked to the *proportion dominance bias*. Third, because demographic data are very common, and should be easily understood by everyone. We used data at two different times steps to ensure the blame for the situations would be put on the Syrian conflict, not on the affected populations. We expected this would leverage the *justice bias*, and that it would ensure some kind of emotional response, as the differences are truly shocking.

A Design Space for Anthropographics

We created eleven anthropographic designs and a baseline pie chart for each story using the data described above. These forty four anthropographics were the basis for our design space. Although we never intended to test all of them—our goal was to assess whether anthropographics generally have an effect on empathy and donating behavior, not to test for most effective designs—and although we do not claim they are exhaustive, we wanted to get a sense of the creative possibilities. Establishing this design space also allowed us to select the most "representative" designs for our experiments (as done in [15]). We found inspiration in [2, 16, 45, 46], and varied our designs according to: **class of visualization**, **human shape**, **unit labelling**, and **unit grouping**.

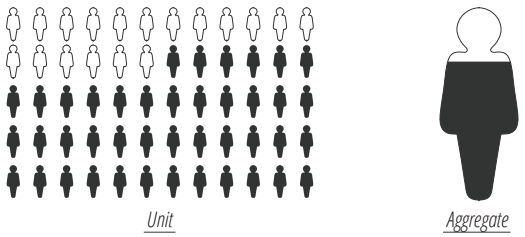

Figure 1. Two classes of visualization: *unit* and *aggregate*.

Class of visualization: we created *unit* and *aggregate* visualizations (Fig. 1). Unit visualizations show each row of a data table using a single visual mark, or *unit* [23], while aggregate visualizations show statistics. Although most anthropographics are unit visualizations (see [1]), and while anthropomorphic aggregation is somewhat of an oxymoron, we found examples of aggregate designs in humanitarian reports (*e.g.,* [3]).

Human shape: we designed a series of pictograms, varying in *realism* and *expressiveness*, to apply to both unit and aggregate visualizations. According to Waytz *et al.*, anthropomorphization depends on the attribution of humanlike physical features, like a face and hands; and of humanlike mental capabilities, like intentions or emotions [65]. The prior can be mapped to the *realism* of the pictogram, while the latter to it's *expressiveness*. This creates a simple $2 * 2$ space (Fig. 2), which can be segmented for simplicity along the *realism* axis into groups

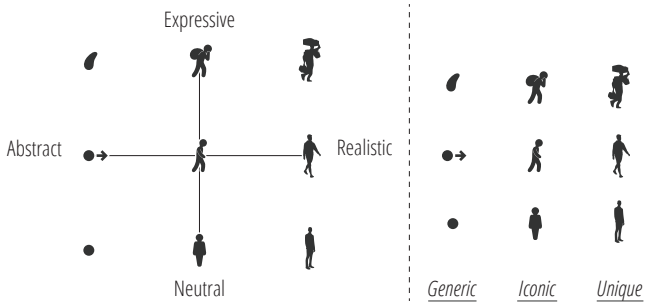

Figure 2. Depicting a migrant: from *abstract* to *realistic*, and from *neutral* to *expressive*; grouped into *generic*, *iconic*, and *unique* pictograms.

of *generic*, *iconic*, and *unique* pictograms, according to their congruence with a unique individual.

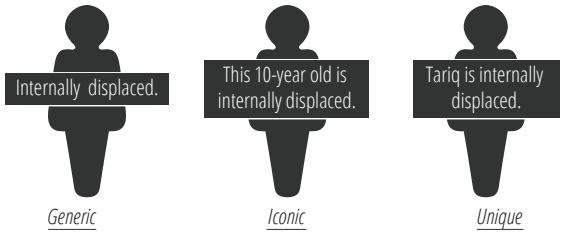

Figure 3. Three ways of labelling units: *generic*, *iconic*, and *unique*.

Unit labelling: we also created *generic*, *iconic*, and *unique* labels (Fig. 3) to make *abstract* and *neutral* pictograms more evocative for first-time viewers [34]. Generic labels provide no personal information, and no sense of individualness. Iconic labels provide some demographic information, and a general sense of individualness. Unique labels provide names, and true individualness. In our experiments, we used popular Arab first names as proxies for real names, since we chose to avoid re-identifiable data for ecologic validity.

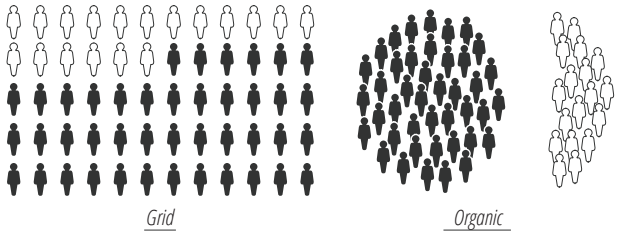

Figure 4. Two ways of grouping units: *grid* and *organic*.

Unit grouping: we created *grid* and *organic* groupings for the unit visualizations (Fig. 4). Grid groupings are visually clear, but they are somewhat unrealistic regarding the way people form groups in real life. Organic groupings are less clear, but more realistic. Ensuring the perception of distinct groups is important, as Bartels & Burnett have found that *group construal*, *i.e.,* the way multiple entities are perceived as either a number of individuals or a single group, can have an effect on the *proportion dominance bias* [5].

Measurements

We measured participants' empathic concern, personal distress, perception of story-protagonists' responsibility, and perception of justice of donating for each HR narrative on 7-point scales. We also collected qualitative feedback through free-from text inputs. For empathic concern and personal distress, we used Batson's list of emotion adjectives [7]. We asked participants to report how strongly they felt *sympathetic*, *moved*, *compassionate*, *tender*, *warm*, and *softhearted* for empathic concern; and *alarmed*, *grieved*, *upset*, *worried*, *disturbed*, *perturbed*, *distressed*, and *troubled* for personal distress. We randomized the order of appearance of these adjectives for each participant. For responsibility, justice, and donation likelihood, we used a series of questions extracted from [43]. We also asked participants a topic availability question to check whether they knew about each topic beforehand. We refer to all these measurements as the *empathy and biases scales*.

We then evaluated prosocial behavior through donation likelihood (on a 7-point scale), and a fictional donation allocation procedure. This provided a realistic scenario for the HR narratives, as most organizations conduct fundraising campaigns to support populations in distress, or to advocate for their rights. We told participants that "someone" was willing to make a $10 donation in their name, and that they could allocate the money as they liked. We believed this would amplify the *moral bias*, since participants would not have to consider donating their own money. We initially asked them to determine how they would split the $10 between stories, but we later changed this to force a dichotomous choice. Finally, we asked participants to briefly explain their choice.

General Procedure

We conducted our experiments on Amazon Mechanical Turk (AMT). Upon accepting the Human Intelligence Task (HIT), and providing consent through AMT, participants were directed to an external webpage, which welcomed them with a short introduction. A single button allowed them start the first narrative. After reading it, they were asked to fill out a first set of empathy and biases scales, and to indicate their donation likelihood. They then moved on to the second narrative, after which they were asked to fill out another set of scales. In the end, they were given the possibility to read through any of the two narratives again, and as many times as they liked. They were also prompted to make the donation allocation. Once the HIT was completed, they were paid $.30.

PRE-TESTING THE STORIES

In this section, we present a pre-study we conducted to select the pair of least diverging topics for our two narratives (similarly to what was done in [50]). These had to be different, but at the same time provoke similar affective responses when presented using the same baseline chart condition, so as to minimize possible *story-topic-related* confounds when comparing different chart conditions in our experiments.

Design Specificities

Our pre-study consisted of six pairwise comparisons of each of our four stories. We used the general materials and methods described in the previous section, but we restricted the chart conditions to the *same* standard pie chart (no anthropographic). We also used the 'split' donation allocation procedure, as we believed it would yield finer-grained results.

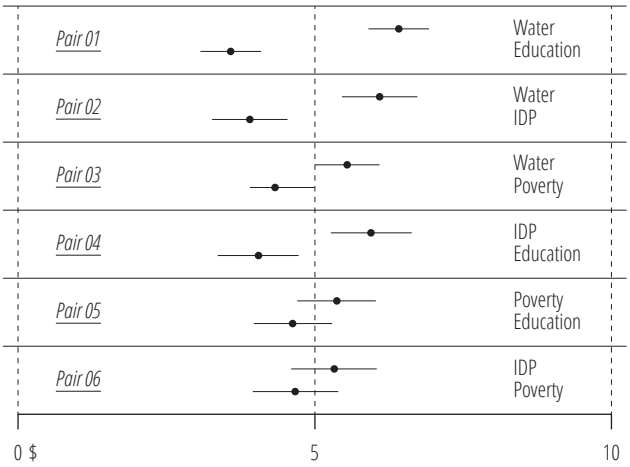

Figure 5. Mean donation allocations for each pairwise comparison.

For each pairwise comparison, we recruited 50 participants located in the United States who were required to have a 99% acceptance rate on AMT. We rejected the data of 16, due to incomplete or irrelevant responses (*e.g.,* participants who simply answered *e.g.,* "brilliant" when asked to describe their impressions about the stories). This resulted in a total of **284** participants (♂ = 155, ♀ = 129).

Results

All the analyses and discussions in this article are based on estimation, *i.e.,* point estimates with 95% confidence intervals (95% CI). The 95% CI are based on 10,000 percentile bootstrap replicates of the parameter of interest. This complies with the recommendations put forward in [4, 20, 22].

We first inspected the differences in donation allocations (Fig. 5), as this should be the most immediate indicator for diverging topics. We then compared aggregated responses to the empathy and biases scales for each story (Fig. 6).

Discussion

Results of the donation allocations clearly show that **water** was the most divergent topic. There is good evidence in Fig. 5 that participants were willing to allocate more money to support access to clean water than to any other cause. **Water** was indeed considered a more vital need, as participants typically explained that "having water is more crucial right now because it is essential to survive," "no one can survive without water," or "the water crisis is more likely to cause deaths and is therefore the more immediate need." This illustrates how rational participants' decisions to donate generally were. We expected this might be different in our experiments, since the anthropographics should elicit more emotional responses.

That said, empathic concern and personal distress were generally high for each story (> 5, see Fig. 6); the perception of responsibility was low, and the perception of justice very high. This simply indicates that: 1) HR narratives about suffering

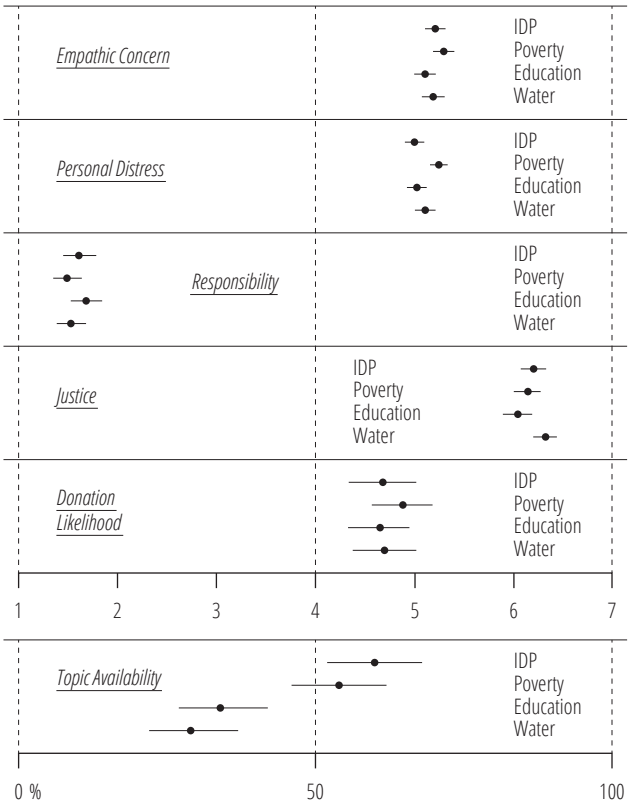

Figure 6. Aggregated mean responses to the empathy and biases scales.

children do indeed elicit empathy; 2) children are not perceived as responsible for their plight; and 3) people think it is just to help these children with donations. Donation likelihood however, was intriguingly consistent, considering how much more money people ended up allocating to the **water** story. This discrepancy leads us to believe that donation likelihood was probably more affected by the *moral bias* [43], *i.e.,* by individual differences in the perceived importance of fairness or justice [57], than the actual allocations, which were more objective and rational. This could explain the lower reliability of the donation likelihood results (shown by the wider error bars [41]). Meanwhile, topic availability was clearly separated into two groups: **IDP** and **poverty** were well known topics, while **education** and **water** were much less so.

Based on these results, we selected the **IDP** and **poverty** stories for our experiments. We excluded the **water** story because of its highly divergent effect on donation allocations. We then decided to leave out the **education** story because the topic was much less known than the other two. Although it is possible that higher topic-awareness may reduce the effect of presentation (*e.g.,* of chart design) on empathy—as people may already have other pre-stored (or *available*) empathy-arousing images in mind—we preferred to use the two stories that had the most similar levels of topic availability.

TESTING THREE ANTHROPOGRAPHICS

In this section, we describe our three initial experiments, in which we compared different anthropographic designs with the pie chart baseline. We used the stories selected in the pre-study, and successively tested the three designs we considered most representative of the eleven we created. Our hypotheses were based on the anthropographic assumption:

- **H1:** the anthropographic should elicit more empathy than the standard pie chart (regardless of the story it is set in) [36, 42, 55, 66]; and therefore
- **H2:** the anthropographic should encourage more prosocial behavior [30, 43, 64], which should be more emotionally-driven than in our pre-study.

While piloting these experiments, we realized that the 'split' donation allocation procedure often led to even distributions ($5 for each story). This is consistent with the findings in [64], which show that empathic concern usually drives people to split resources between multiple charities, rather than giving everything to a single charity. While such fine-grain results were useful in our pre-study, and although they suggested early on that we might fail to confirm **H2**, we wanted to encourage more emotionally-driven decisions. Therefore, we changed the procedure to force a dichotomous choice: participants would have to allocate all $10 to a single story.

We also noticed that very few participants hovered over the units to display the labels. This meant they did not see them. As labels were an important attribute of our anthropographic designs, we added a Suggested Interactivity cue [15] that sequentially displayed the labels of random units. That way, participants would see the labels, even without interaction.

Experiment 1

In our first experiment, we tested a *unit* anthropographic organized in *organic* groupings, in which units had *unique* shapes and labels (the **unique individuals design**). We chose to test it first, as we believed the uniqueness and higher realism of the units would most effectively tap into participants' experiential knowledge of postures and attitudes [49, 52], thus facilitating empathy-arousal through mimicry and direct association [38], and ultimately helping to reproduce the *persona* and/or *identifiable victim effect* (as did the "fully expressive" character in [44]). We expected this would elicit higher levels of empathy, and therefore would attract more donations (**H1** and **H2**). We used our general materials and methods, and included the minor adjustments mentioned above. We recruited 50 participants on AMT who had not participated in our pre-study, setting the same requirements. We rejected the data of 2 who performed the HIT twice. This resulted in a subset of **48** participants (♂ = 23, ♀ =25).

Results

This time, we first inspected responses to the empathy scales between chart conditions (Fig.7). We then checked for possible interactions between empathy and participant variables. We separated participants by gender, then by topic availability (Fig.8). We expected people without prior knowledge of the topics might be more affected by the mode of presentation.

We then looked at donation likelihood between chart conditions (Fig. 9). Here too, we checked for possible interactions with participant variables. Although we do not show the results for responsibility and justice here due to space limitations,

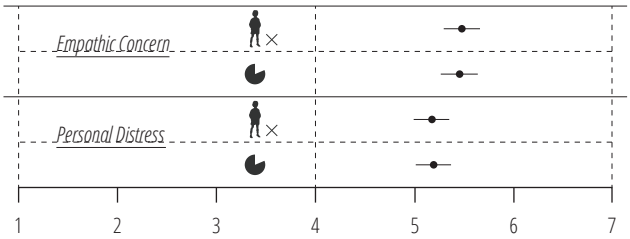

Figure 7. Mean responses to the empathic concern and personal distress adjectives between chart conditions. For each emotional component, mean responses in the anthropomorphized data graphic condition are shown on top; those for the pie chart baseline are shown below.

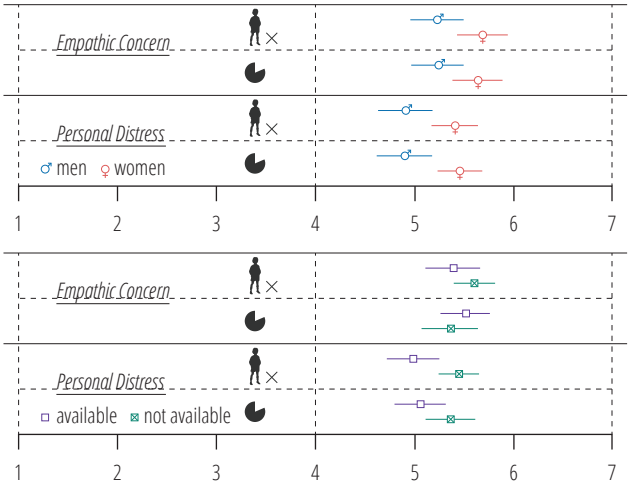

Figure 8. Mean responses separated by gender and topic availability.

these were generally consistent with those of our pre-study. As we believe it is clear people generally perceive that children are not responsible for their plight (*responsibility bias*), and that it is just to help them (*justice bias*), we do not mention these possible biases any further. Finally, we inspected the results of the dichotomous donation allocations. **43.7%**, 95% CI [30.7%, 57.7%] of participants allocated the $10 to the narratives using the anthropographic.

Discussion

Our results do not support **H1**. Fig. 7 shows that the **unique individuals design** and the standard pie chart had very similar effects on viewer's empathic concern and personal distress. The results are highly reliable (short error bars), and the fact that there is some difference between these emotional components (independently of chart condition) is a good indication that participants did not respond randomly—as the list of adjectives we used to assess both feelings was randomly ordered for each participant. In addition, while Fig. 8 shows no evidence of an interaction between gender, chart type, and empathy, it shows that women were generally more empathic than men. This is consistent with [26]. Fig. 8 also shows no real evidence of an interaction between topic availability, chart type, and empathy. The only slightly perceptible effect of the **unique individuals design** is on the difference in personal distress between people who knew the stories beforehand and people who did not. However, it is possible that this is confounded by

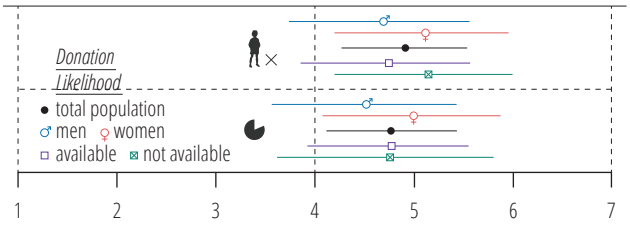

Figure 9. Mean donation likelihood.

other mediating factors like gender or story—as there was already evidence in Fig. 6 of a very slight difference in personal distress between the **IDP** and **poverty** stories.

H2 is not supported either. Fig. 9 shows no evidence of a difference in donation likelihood between chart conditions, whatever the participant subgroup. That said, the reliability of these results is quite low (wide error bars), which comforts our idea that there are individual differences. In addition, the **unique individuals design** attracted less donations than the pie chart—although the evidence is weak. Once again, participants explained their donation allocations very rationally, which suggests that neither the anthropographic, nor the forced dichotomous choice truly encouraged them to act on emotions induced by the mode of presentation. Only one participant mentioned that "the first slideshow [which used the **unique individuals design**] was more moving."

Overall, our results suggest that the **unique individuals design** did not reproduce the expected *persona* and/or *identifiable victim effect*. Upon reflection, we considered this may have been due to a *semantic incongruence* between the type of data and their representation: the anthropographic showed abstract statistics, *i.e.,* normalized proportions, as groups of unique individuals, not as "abstract statistics." Although we had not anticipated that this might have an effect, since it is relatively common practice (see [1]), participants may have considered it deceptive [51], and their feelings may have been confounded.

Experiments 2 and 3

In our second and third experiments, we attempted to reduce the semantic incongruence of the **unique individuals design** by lowering the uniqueness of each anthropographic. We expected this might positively impact participants' empathy and prosocial behavior. In Exp. 2, we tested a *unit* anthropographic organized in *grid* groupings, in which units had *iconic* shapes and *generic* labels (the **iconic individuals design**). In Exp. 3, we tested an *aggregate* anthropographic with an *iconic* shape (the **iconic statistic design**). Although the efficiency of this latter design is doubtful from a visual perception perspective, it had the advantage of being the least semantically incongruent of all our designs. For both experiments, we used the same materials, methods, and hypotheses as in Exp. 1. We recruited 50 new participants for each, and retained the work of **46** in Exp. 2 (♂ = 19, ♀ = 27), and of **45** in Exp. 3 (♂ = 19, ♀ = 26).

Results

We conducted exactly the same analysis as in Exp. 1. All results for the empathy scales and donation likelihood for Exps. 2 and 3 are respectively shown in Figs. 10 and 11. **34.8%**, 95%

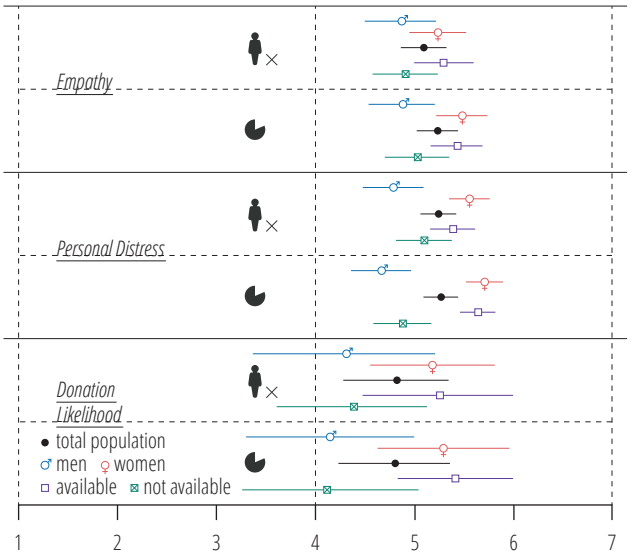

Figure 10. Results for Exp. 2: the iconic individuals design.

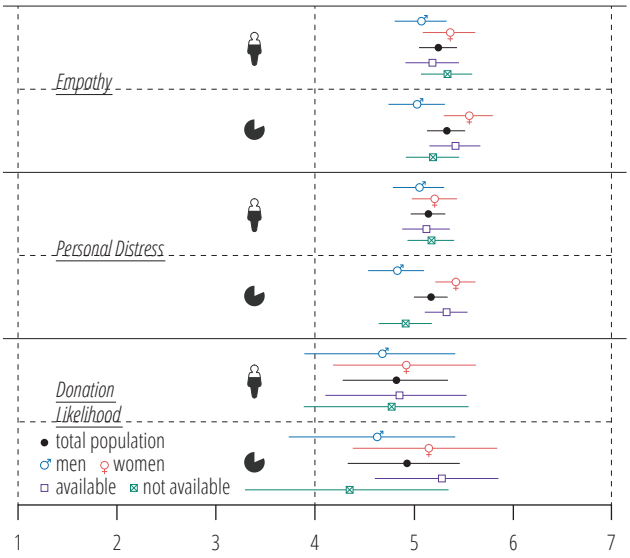

Figure 11. Results for Exp. 3: the iconic statistic design.

CI [22.6%, 49.2%] of participants in Exp. 2, and **51.1%**, 95% CI [37%, 65%] in Exp. 3 allocated the $10 to the narratives using the anthropographics.

Discussion

Although we see a higher variability *within* participant subgroups and chart conditions (Fig. 10), our results generally do not support **H1** in either experiment. The anthropographics and the standard pie chart had similar effects on empathic concern and personal distress. In fact, our results even show very slight negative trends. We believe these are likely the result of individual differences in empathic disposition [24]—people may be more or less empathic in general. There is also no real evidence of interactions between gender or topic availability, and chart type and empathy. That said, once again women were generally more empathic than men, and the effect of availability seems confounded by other mediating factors.

H2 is not supported either. Donation likelihood is generally consistent between chart conditions in both experiments—even though the reliability of these results is still low. In addition, there is some evidence that the **iconic individuals design** actually attracted less donations than the pie chart; and there is no real evidence that the **iconic statistic design** attracted more. Explanations were again very rational, even if people reason differently, *e.g.,* "I wanted to make sure the children could find homes. I figure this could also help children living in poverty so it's a win win," or "I think that giving them homes is more of a short term measure. Permanently getting these children out of poverty would be best for the long term."

Initial Experiments' Discussion

These initial results fail to validate the anthropographic assumption. Empathic concern, personal distress, and donation likelihood were very similar between chart conditions, and across experiments. Donation allocations however, were somewhat inconsistent. Even our prime candidate, the **unique individuals design**, seems to have failed to reproduce the *persona* and/or *identifiable victim effect*. Although we had speculated on the possible confounding effect of a semantic incongruence after Exp. 1, Exps. 2 and 3 consistently failed to show a difference when trying to reduce that incongruence. Upon further reflection, this failure to reproduce *persona* and/or *identifiable victim effect* may be because (at least) the latter is assumed to be linked to face recognition mechanisms in the brain [30] (the human shapes in our designs had no faces). Another possible confound could be that our choice of unit labeling in the **unique individuals design** suffered from the *familiarity bias* [38]. The *unique* labels showed Arab names, so our American participants may have perceived the children in the stories as out-group members, which could have limited the effect of the anthropographic. Nevertheless, empathy levels were high in all three experiments (> 5), which leads us to believe the baseline level of empathy elicited by our HR narratives (independent of chart conditions) may have reached a threshold. This would suggest that the anthropographics showed no effect simply because there was no room for that effect to be visible.

EXTENSIONS

In this section, we present four follow-up experiments, in which we adjusted different aspects of our general materials and methods, in an attempt to lower the baseline level of empathy elicited by our HR narratives. Although this made the narratives somewhat less ecologically valid, we hoped it would accentuate the effect (if any) of the anthropographics.

Design Specificities

As discussed at the end of our pre-study, the high awareness of the **IDP** and **poverty** topics may have reduced the effect of presentation on empathy. In Exp. 4, we replaced the stories we had previously used with the **water** and **education** stories. The standard pie chart may have elicited unforeseen affective responses, which may have emphasized its effect on empathy. In Exp. 5, we compared the pie chart with *text alone*. The fact that the varying chart conditions were only a relatively small

part of the data-driven stories (2/5 slides) may have favored the text as a reminder of the human dimension of the data. In Exp. 6, we stripped the stories to keep only the slides with the varying charts conditions (slides 3 and 4). Finally, Borkin *et al.* have found that people are generally more attracted to textual elements in visualizations, and that *e.g.,* titles have a significant effect on what they take away [13]. The wording of the text, and the focus on children may have been so emotionally charged that they overshadowed the effects of the anthropographics. In Exp. 7, we again restricted the stories to slides 3 and 4, and we removed all references to children, and re-edited the text to make it shorter and dryer.

Considering the *familiarity bias*, we created a modified version of the **unique individuals design** for Exps. 4 and 6: the **familiar individuals design**. We replaced the *unique* labels with *iconic* labels, changing the Arab names to age identifiers. To comply with our re-editing of the text in Exp. 7, we also created the **generic individuals design**, using *generic* labels. We kept the rest of the materials, methods, and hypotheses the same as before—with a slight exception for Exp. 5, in which, for consistency, we hypothesized that the pie chart should elicit more empathy and encourage more prosocial behavior than text alone. We recruited 50 new participants for each experiment, and retained the work of **49** in Exp. 4 (♂ = 22, ♀ = 27); of **40** in Exp. 5 (♂ = 24, ♀ = 16); of **47** in Exp. 6 (♂ = 25, ♀ = 22); and of **47** in Exp. 7 (♂ = 21, ♀ = 26).

Results

We first compared the aggregated levels of empathy for each follow-up experiment with those of our initial experiments (Fig. 12) to determine whether our adjustments had an effect on the overall level of empathy—and thereby on the baseline level of empathy. We then inspected responses to the empathy scales and donation likelihood between chart conditions. Due to space limitations, we do not detail the results for Exps. 4 and 5 here. Fig. 12 shows that the adjustments we made in these experiments failed to lower the baseline level of empathy, and the different chart conditions once again showed no evidence of differences[1]. Results for Exps. 6 and 7 however, are respectively shown in Figs. 13 and 14, as the adjustments in these clearly lowered the baseline (Fig. 12). For simplicity, we do not show the gender differences or topic availability here. We have found that the prior simply highlights a general *gender bias* in empathy-arousal [26], and that the latter is likely confounded by other mediating factors. Finally, we inspected the results of the donation allocations. **51%**, 95% CI [37.2%, 64.7%] of participants in Exp. 6, and **61.7%**, 95% CI [47.4%, 74.2%] in Exp. 7 allocated the $10 to the narratives using the anthropographics.

Discussion

Despite our adjustments, Exps. 4 and 5 clearly failed to lower the baseline level of empathy (> 5, see Fig. 12). This suggests that neither the topics, nor the standard pie chart were responsible. In addition, empathic concern, personal distress, and donation likelihood in these experiments were once again very similar between chart conditions; and the donation allocations showed no real evidence of an effect of chart conditions.

[1] Results for Exps. 4 and 5 can be found in the supplemental material—along with those of all other experiments.

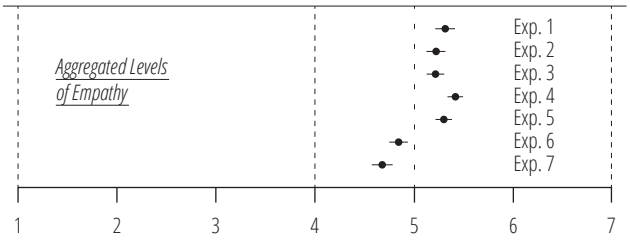

Figure 12. Aggregated levels of empathy for each experiment.

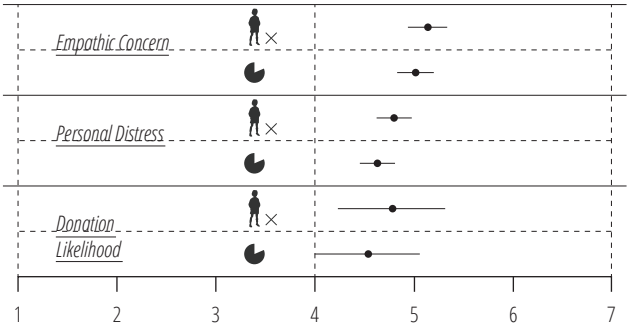

Figure 13. Results for Exp. 6: stripped stories.

Exps. 6 and 7 however, did lower the baseline (< 5). This suggests that the extra contextual information and the wording were likely responsible for the higher levels of empathy. While we expected lowering the baseline would accentuate the difference between chart conditions, Figs. 13 and 14 show it did not. Although there is a very slight trend in favor of the anthropographic in Fig. 13, the evidence is weak, and it is not confirmed in Fig. 14—where the aggregated level of empathy is lowest. Likewise, donation allocations show that the anthropographics attracted slightly more money than the pie chart in both experiments, but the evidence is not very strong. In addition, we believe all the results of the donation allocation procedure should be considered with caution, as they may have been confounded by an anthropomorphic carryover effect. Participants may have automatically transferred the human dimension of the data from the narrative with the anthropographic to the one without, as donation decisions were made at the end, after having read through both narratives.

Thus, once more our results generally do not support **H1** and **H2**. We conclude that a tragic HR narrative simply *is* a tragic HR narrative. Removing content and making the text dryer may affect readers' empathy, but presenting it with anthropographics instead of standard charts makes no real difference: empathic readers will inevitably feel *for* the suffering protagonists. This generally complies with the concept of narrative empathy [40], and indicates that when the graphics are part of an emotive narrative, they can afford to be abstract.

GENERAL DISCUSSION

All our results invalidate the anthropographic assumption, at least in the specific narrative context in which we tested it. Anthropographics and standard charts had very similar effects

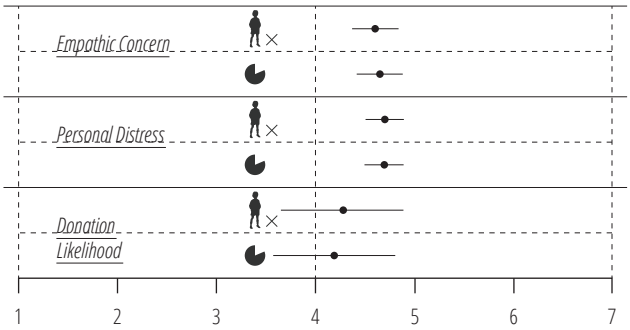

Figure 14. Results for Exp. 7: stripped and reworded stories.

on participants' empathy, as well as on their prosocial behavior in all our experiments. Although it may be argued that the emphasis we put on being ecologically valid (in the structure, content, and wording of the HR narratives) may have initially confounded our results, Figs. 13 and 14 suggest it did not. Lowering the impact of other materials did not promote the anthropographics, which leads us to believe that when narrative empathy [40] is high, charts can afford to be abstract.

Participants generally reported a high donation likelihood (> 5). Interestingly though, as their empathy decreased, so did this likelihood. This concurs with the idea that donation likelihood is mediated by empathy [43, 64]. However, when it came to allocating donations, their decisions were mostly rational. Although we expected the anthropographics would encourage more emotionally-driven decisions, we failed to find this true. We relate this to the results of [30], which show that abstract silhouettes and names have a lesser effect than photos on viewers' tendency to donate to victims.

More generally, our results concur with a number of previous findings. In addition to the mediating role of empathy on donation likelihood [43, 64], we found that women are generally more empathic than men [26]; that empathic concern and personal distress are related [7]; and that there are likely some individual differences in dispositional empathy [24]. This provides strong external validity, while the overall consistency of our results provides good internal validity.

However, although our results do not show a higher positive effect of anthropographics on empathy and prosocial behavior compared to standard charts, they do not show a negative effect either. In other words, although anthropographics were not truly beneficial in our experiments, they were not detrimental. We stress this, because in more realistic settings where people may not be as biased toward information as in a controlled experiment, anthropographics may be useful for *e.g.,* attracting readers' attention, and for drawing them into the narrative [35]. This might ultimately be useful for eliciting empathy. As such, we consider our results an initial assessment of the empathic and prosocial power of anthropographics. They complement Bateman *et al.*'s call to learn more about the effects of different types of visual embellishments in charts [6], and open a number of exciting avenues for future work on the anthropographic assumption. Typically, focusing on other subjects than the highly mediatized Syrian crisis may yield different results. Expanding our design space for anthropographics to more detailed drawings with *e.g.,* faces (like in [18, 45]), or even to photos (when possible) may facilitate reproducing the *persona* and/or *identifiable victim effect*, whereby the more lifelike renderings should have a stronger positive effect on viewers' affect [44]—although they may also increase the semantic incongruence between graphics and data if the prior do not map unique human shapes to unique individuals. Animating the human shapes may help too. Using more subtle cues like visual metaphors [12] (as done in [53]) or *motionscapes* [29] may tap deeper into viewers' experiential knowledge [49, 52], which may further facilitate empathy-arousal through mimicry and direct association [38]. Adding cinematographic and sound effects, like the dramatic zooming-in-and-out and powerful voice-over narration used in [33] might also increase affective responses. In contexts where disaggregated data may be more readily usable, lower data-granularity may contribute to further reducing the semantic incongruence, which may positively impact viewers' perception of the uniqueness of units. Similarly, increasing the amount of units to show absolute values (instead of normalized values) may provide more *concrete scales* [19], *i.e.,* scales that are easier to relate to, which may also help viewers better grasp the magnitude of certain HR tragedies (as done in [61]). Finally, further tweaking our general materials and methods may lead to finer results, especially regarding donating behavior. A different, between-subjects procedure may remove the possible confounds of the donation allocation, and changing the fictional allocation to a real donation (similarly to what is done in [30]) may incite participants to think differently about how they allocate the money.

CONCLUSION AND FUTURE WORK

In this article, we have presented a design space for anthropographics, and a series of experiments, in which we compared various anthropographic designs set in human rights narratives with a standard chart, to investigate the growing assumption that because anthropomorphized data graphics create an immediate visual connection between abstract data and actual people [36], they are better suited than standard charts for eliciting empathy and prosocial behavior [42, 66]. Contrary to this assumption and our expectations, we have found that both types of visualizations have very similar effects. This initial assessment suggests that when the graphics are part of a broader emotive narrative, they can afford to be abstract.

However, while this seems to imply that anthropographics are not truly beneficial for eliciting empathy and prosocial behavior, we stress it does not mean they are detrimental. We strongly believe our results should not discourage further exploration and use of such graphics. We personally intend to extend our design space, and we have proposed a number of possible avenues for future work, which, although we acknowledge are not exhaustive, contribute to the elaboration of a research agenda for anthropographics, and ultimately to an agenda for better understanding the effects of different types of visual embellishments in information visualization.

ACKNOWLEDGEMENTS

This research was supported by a John D. And Catherine T. MacArthur Foundation grant.

REFERENCES

1. A Collection of anthropographics. http://myjyby.github.io/Athropographics/www/.
2. A Hommage to Gerd Arntz. http://www.gerdarntz.org/.
3. World Humanitarian Data and Trends. http://online.fliphtml5.com/srhi/mtgf/#p=35, 2015.
4. American Psychological Association. *The Publication manual of the American psychological association (6th ed.)*. Washington, DC, 2010.
5. D. M. Bartels and R. C. Burnett. A group construal account of drop-in-the-bucket thinking in policy preference and moral judgment. *Journal of Experimental Social Psychology*, 47(1):50–57, 2011.
6. S. Bateman, R. L. Mandryk, C. Gutwin, A. Genest, D. McDine, and C. Brooks. Useful Junk?: The Effects of Visual Embellishment on Comprehension and Memorability of Charts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2573–2582, New York, NY, USA, 2010. ACM.
7. C. D. Batson. Prosocial Motivation: Is it ever Truly Altruistic? volume 20 of *Advances in Experimental Social Psychology*, pages 65–122. Academic Press, 1987.
8. C. D. Batson. *The Altruism Question: Toward a Social Psychological Answer*. L. Erlbaum, 1991.
9. C. D. Batson. Empathy-Induced Altruistic Motivation. *Prosocial motives, emotions, and behavior: The better angels of our nature*, pages 15–34, 2010.
10. D. Bischof-Köhler. The development of empathy in infants. In *Infant Development: Perspectives from German Speaking Countries*, chapter 12, pages 245–273. Lawrence Erlbaum Associates, Inc, 1991.
11. P. Bloom. Against Empathy. http://bostonreview.net/forum/paul-bloom-against-empathy, 2014.
12. R. Borgo, A. Abdul-Rahman, F. Mohamed, P. W. Grant, I. Reppa, L. Floridi, and M. Chen. An Empirical Study on Using Visual Embellishments in Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2759–2768, Dec 2012.
13. M. A. Borkin, Z. Bylinskii, N. W. Kim, C. M. Bainbridge, C. S. Yeh, D. Borkin, H. Pfister, and A. Oliva. Beyond Memorability: Visualization Recognition and Recall. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):519–528, Jan 2016.
14. M. A. Borkin, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, and H. Pfister. What Makes a Visualization Memorable? *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2306–2315, 2013.
15. J. Boy, L. Eveillard, F. Detienne, and J. D. Fekete. Suggested Interactivity: Seeking Perceived Affordances for Information Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):639–648, 2016.
16. C. Burke, E. Kindel, and S. Walker. *Isotype: Design and Contexts 1925–1971*. Princeton Architectural Press, 2014.
17. A. Cairo. Empathy and Visualization. http://www.thefunctionalart.com/2016/02/empathy-and-visualization.html, 2016.
18. M. Chalabi. https://www.instagram.com/mona_chalabi/.
19. F. Chevalier, R. Vuillemot, and G. Gali. Using Concrete Scales: A Practical Framework for Effective Visual Depiction of Complex Measures. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2426–2435, 2013.
20. G. Cumming. The New Statistics: Why and How. *Psychological science*, 25(1):7–29, 2014.
21. M. H. Davis. Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1):113–126, 1983.
22. P. Dragicevic. Fair Statistical Communication in HCI. In *Modern Statistical Methods for HCI*, pages 291–330. Springer, 2016.
23. S. M. Drucker and R. Fernandez. A Unifying Framework for Animated and Interactive Unit Visualizations. Technical Report MSR-TR-2015-65, 2015.
24. C. J. Einolf. Empathic concern and prosocial behaviors: A test of experimental results using survey data. *Social Science Research*, 37, 2008.
25. N. Eisenberg. Empathy-Related Responding: Links with Self-Regulation, Moral Judgment, and Moral Behavior. *Prosocial motives, emotions, and behavior: The better angels of our nature*, pages 129–148, 2010.
26. N. Eisenberg, A. Cumberland, I. K. Guthrie, B. C. Murphy, and S. A. Shepard. Age Changes in Prosocial Responding and Moral Reasoning in Adolescence and Early Adulthood. *Journal of Research on Adolescence*, 15(3):235–260, 2005.
27. N. Eisenberg and P. A. Miller. The Relation of Empathy to Prosocial and Related Behaviors. *Psychological Bulletin*, 101(1):91–119, 1987.
28. J. Emerson. Ten Challenges to the Use of Data Visualization in Human Rights. http://backspace.com/notes/2016/02/ten-challenges.php, 2016.
29. B. L. G. D. Feng, C. Beyond Data: Abstract Motionscape as Affective Visualization. *Leonardo/ISAST*, 2014.
30. A. Genevsky, D. Västfjäll, P. Slovic, and B. Knutson. Neural Underpinnings of the Identifiable Victim Effect: Affect Shifts Preferences for Giving. *The Journal of Neuroscience*, 33(43):17188–17196, 2013.

31. C. C. Gould. *Globalizing Democracy and Human Rights*. Cambridge University Press, 2006.
32. T. Grodal. Visual Communication as Textsigns and as Embodied Mental Processes. In *Nordic Visual Communication Conference, Oslo*, 1994.
33. N. Halloran. The Fallen of World War II. http://www.fallen.io/ww2/.
34. K. Haramundanis. Why icons cannot stand alone. *SIGDOC Asterisk J. Comput. Doc.*, 20(2):1–8, May 1996.
35. S. Haroz, R. Kosara, and S. L. Franconeri. Isotype visualization: Working memory, performance, and engagement with pictographs. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 1191–1200, New York, NY, USA, 2015. ACM.
36. J. Harris. Connecting with the Dots. https://source.opennews.org/en-US/learning/connecting-dots/, 2015.
37. M. A. Harrison and A. Hall. Anthropomorphism, Empathy, and Perceived Communicative Ability Vary with Phylogenetic Relatedness to Humans. *Journal of Social, Evolutionary, and Cultural Psychology*, 4(1), 2010.
38. M. Hoffman. Prosocial Behavior and Empathy: Developmental Processes. In N. J. S. B. Baltes, editor, *International Encyclopedia of the Social & Behavioral Sciences*, pages 12230–12233. Pergamon, Oxford, 2001.
39. K. Jenni and G. Loewenstein. Explaining the Identifiable Victim Effect. *Journal of Risk and Uncertainty*, 14(3):235–257, 1997.
40. S. Keen. Narrative Empathy. http://wikis.sub.uni-hamburg.de/lhn/index.php/Narrative_Empathy, 2013.
41. M. Krzywinski and N. Altman. Points of Significance: Error bars. *Nature Methods*, 10:921–922, 2013.
42. S. Lambert. And What Do I Do Now? Using Data Visualization for Social Change. http://artisticactivism.org/2016/01/data-visualization-for-what/, 2016.
43. S. Lee, K. P. Winterich, and W. T. Ross. I'm moral, but I won't help you: The distinct roles of empathy and justice in donations. *Journal of Consumer Research*, 41(3):678–696, 2014.
44. J. C. Lester, S. A. Converse, S. E. Kahler, S. T. Barlow, B. A. Stone, and R. S. Bhogal. The persona effect: Affective impact of animated pedagogical agents. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, pages 359–366, New York, NY, USA, 1997. ACM.
45. S. McCloud. *Understanding Comics*. A Kitchen Sink book. HarperCollins, 1994.
46. S. McCloud. *Making Comics*. HarperCollins, 2011.
47. N. M. McDonald and D. S. Messinger. The Development of Empathy: How, When, and Why. In A. A., L. J. A, and S. J. J., editors, *Free Will, Emotions, and Moral Actions: Philosophy and Neuroscience in Dialogue*, pages 341–368. IF-Press Morolo, 2011.
48. J. Miksatko, K. H. Kipp, and M. Kipp. The persona zero-effect: Evaluating virtual character benefits on a learning task with repeated interactions. In *International Conference on Intelligent Virtual Agents*, pages 475–481. Springer, 2010.
49. S. E. Moriarty. Abduction: A Theory of Visual Interpretation. *Communication Theory*, 6(2):167–187, 1996.
50. A. V. Pandey, A. Manivannan, O. Nov, M. Satterthwaite, and E. Bertini. The persuasive power of data visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):2211–2220, 2014.
51. A. V. Pandey, K. Rall, M. L. Satterthwaite, O. Nov, and E. Bertini. How Deceptive Are Deceptive Visualizations?: An Empirical Analysis of Common Distortion Techniques. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 1469–1478, 2015.
52. C. S. Peirce. *Peirce on signs*.
53. Periscopic. U.S. Gun Deaths. http://guns.periscopic.com/?year=2013.
54. L. D. Riek, T.-C. Rabinowitch, B. Chakrabarti, and P. Robinson. How Anthropomorphism Affects Empathy Toward Robots. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction*, HRI '09, pages 245–246, 2009.
55. J. Schwabish. Policyviz Podcast Episode #31: Rees & Mushon on DataViz Empathy. http://policyviz.com/rees-mushon-on-dataviz-empathy/, 2016.
56. E. Segel and J. Heer. Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, 2010.
57. J. Shang, A. R. II, and R. Croson. Identity Congruency Effects on Donations. *Journal of Marketing Research*, 45(3):351–361, 2008.
58. S. Slobin. What if the Data Visualization is Actually People? https://source.opennews.org/en-US/learning/what-if-data-visualization-actually-people/, 2014.
59. P. Slovic, M. L. Finucane, E. Peters, and D. G. MacGregor. The Affect Heuristic. *European Journal of Operational Research*, 177(3):1333–1352, 2007.
60. P. Slovic and D. Zionts. Can International Law Stop Genocide when our Moral Intuitions Fail Us? In D. J. Ryan Goodman and A. K. Woods, editors, *Understanding Social Action, Promoting Human Rights*, pages 100–128. 2012.

61. The New York Times. Death in Syria. http://www.nytimes.com/interactive/2015/09/14/world/middleeast/syria-war-deaths.html?action=click&contentCollection=Middle%20East&region=Footer&module=WhatsNext&version=WhatsNext&contentID=WhatsNext&moduleDetail=undefined&pgtype=Multimedia&_r=0.

62. E. R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT, USA, 1986.

63. A. Tversky and D. Kahneman. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2):207–232, 1973.

64. G. A. Verhaert and D. V. den Poel. Empathy as added value in predicting donation behavior. *Journal of Business Research*, 64(12):1288–1295, 2011.

65. A. Waytz, J. Cacioppo, and N. Epley. Who Sees Human? The Stability and Importance of Individual Differences in Anthropomorphism. *Perspectives on psychological science : a journal of the Association for Psychological Science*, 5:219–232, 2014.

66. M. Zer-Aviv. DataViz—The UnEmpathic Art. https://responsibledata.io/dataviz-the-unempathetic-art/, 2015.

67. M. Zer-Aviv. Responsible Data Forum: Visualization. https://responsibledata.io/forums/data-visualization/, 2016.