Tell don't just show: Explanatory narrative may improve recall more than exploratory interactivity in communicative visualization

Abhraneel Sarma and Matthew Kay

Abstract— Communicative visualizations incorporating narratives and/or interactivity are commonplace. However, the relative importance of narratives versus interactivity in improving readers' understanding is unclear. As narratives and interactivity are broad design strategies, for the purposes of our work we define more specific subsets of these two strategies: exploratory interactivity (interactivity which allows user exploration of the data) and explanatory narrative (textual messaging with an author–prescribed sequence of steps). We designed visualizations which vary in the presence or absence of exploratory interactivity and explanatory narrative, presented them to Turkers, and measured recall of facts from the visualizations using a ten item True/False questionnaire. We find a weak positive effect—an increase of ~10 percentage points, 95% CI: [6.1, 14.3]—from the presence of explanatory narratives, but likely little or no practical effect from exploratory interactivity (mean: 2.3 percentage points, 95%CI: [–1.4, 6.0]). We argue that explanatory narratives may better facilitate recall than exploratory interactivity in communicative visualizations. Given the expertise required to construct exploratory interactive visualizations, the associated costs (in terms of design time or personnel) may not be worth the likely small gain in reader understanding. We suggest that in contrast to visualization for data analysis (where exploratory interactivity is quite powerful), for communicative visualization, explanatory narrative and other forms of interactivity (e.g., active learning approaches) might represent better trade-offs in the design space.

Index Terms—Communicative Visualization, Narratives, Interactivity

1 Introduction

Communicative visualizations, which have become widely used in data journalism, can have many goals [36, 41]. In this paper, we consider one such goal: conveying a particular message, or knowledge, to the reader through a visualization. To deliver an intended message to the audience, communicative visualizations might employ two high-level design strategies: narrative and interactivity. A narrative presents meaningful information drawn from the data in the form of a story, which can help readers understand data more effectively and efficiently [11]. Interactivity allows performing operations on the data to explore or dig deeper into the data, which can help readers make sense of the data or form their own conclusions from the data [34, 50].

Interactivity has traditionally been considered an essential component of information visualization, including communicative visualizations. Consider Card et al.'s definition of information visualization: "the use of computer supported, **interactive**, visual representations of abstract data to amplify cognition" [18]. Lima's *Information Visualization Manifesto* also emphasizes interactivity, giving it a section unto itself titled "Interactivity is Key" [33]. We suspect this perspective on interactivity stems from the field's traditional focus on supporting exploratory data analysis, where interactivity is vital to data exploration. As our field expands to study communicative visualization, we should re-examine the centrality of interactivity.

Following several years of the prevalence of "interactives" in data journalism, we have noted a shift (at least anecdotally) back towards static designs in communicative visualizations. Outlets such as the New York Times or the Financial Times have started creating visualizations with little or no interactivity [45,47]. Aisch [2] and Tse [47], prominent data journalists, have argued that a large percentage of their readers do not make use of interactive elements in their visualizations [2, 45, 47]. These comments sparked an ongoing debate about the benefits of interactivity in communicative visualizations [3, 10], and raise an important question for designers: *how much does interactivity improve readers' understanding of a communicative visualization?* [1]

Similar questions could be asked of the benefit of narrative. In previous work, authors have claimed that using a narrative in visualizations should allow efficient communication of large amounts of information [21, 49], but these potential benefits have not been investigated. Hence, we ask: *how much does narrative actually improve readers' understanding of a communicative visualization?*

Answering these questions is complicated by the fact that "narrative" and "interactivity" are very broad design strategies. To make headway on these broader questions, we must tear off some piece of "narrative" and "interactivity" to study in particular. To that end, we identified easy-to-operationalise definitions of more specific subsets of these two strategies from the existing literature:

1. *explanatory narrative*: the presence of textual messaging and an author-prescribed sequence of steps through the visualization

2. *exploratory interactivity*: the presence of interactivity that allows the user to explore the data

We cannot (and do not) claim to study "narrative" and "interactivity" in the broad sense; rather, by studying explanatory narrative and exploratory interactivity, we contribute some initial understanding of the shape of this broader space. Armed with our more specific definitions, we systematically vary the the presence or absence of explanatory narrative and exploratory interactivity ($2 \times 2 = 4$ conditions) on the ability of visualization to convey a designer's intended message to a reader.

Operationalising the notion of *a reader's understanding of a designer's intended message* is similarly challenging. We adopt a structured approach based on techniques from the education literature: the revised Bloom's taxonomy [6, 12]. We use this taxonomy to develop questions assessing a specific aspect of learning: recall of factual knowledge. We conducted an online study to evaluate the effects of our four conditions on recall of factual knowledge presented through a visualization. To decrease the likelihood that our results are specific to one visualization design, we modified four different professionally-produced communicative visualizations, yielding 16 different designs

• *Abhraneel Sarma is with the University of Michigan. E-mail: abhsarma@umich.edu.*
• *Matthew Kay is with the University of Michigan. E-mail: mjskay@umich.edu.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

[1] It is worth noting that there may be other benefits to interactivity besides knowledge transfer; e.g. to increase readership through engagement, or allowing skeptical readers to inspect the data more deeply [3]. In this paper we concentrate on how interactivity affects the ability to convey a specific message.

(presence/absence of explanatory narrative × presence/absence of exploratory interactivity × 4 base visualizations).

In a pre-registered, primary analysis, we measured the effects of the presence or absence of explanatory narrative and exploratory interactivity on recall (the probability of correctly answering a question on a ten-item True/False questionnaire). We find:

- The presence of *explanatory narrative* resulted in higher recall. A participant in the explanatory narrative condition was, on average, 10 percentage points (95% CI: [6.1, 14.3]) more likely to answer a question correctly. This effect is equivalent to answering one more question correctly on average on the ten item questionnaire.

- The presence of *exploratory interactivity* likely has little practical effect on recall. A participant in the exploratory interactivity condition was, on average, 2.3 percentage points (95% CI: [-1.4, 6.0]) more likely to answer a question correctly. The upper end of this interval represents answering only about half a question more correctly on average.

These results suggest that, for communicative visualizations, if the designer's goal is to convey an intended message to the reader, the presence of explanatory narrative may be more effective at improving readers' recall of a particular message than exploratory interactivity.

While not the primary focus of this work, **we also conducted two exploratory analyses** to investigate whether commonly-used metrics of user behavior, such as time spent and number of interactions performed with a visualization, can be used as proxies to measure how well a designer's intended message was conveyed.

Since implementing interactive visualizations across multiple devices can incur significant design time and labour costs, we posit that including complex, difficult-to-implement exploratory interactivity may not be worth the likely small gains in readers' learning. For journalistic outfits—where storytelling expertise may be more readily available—explanatory narrative may be simultaneously easier to implement and more likely to increase recall. These trade-offs in learning and implementation complexity may help explain recent trends away from complex exploratory "interactives" in everyday data journalism.

That said, while we did not study them here, other forms of interactivity may represent better cost/benefit trade-offs than exploratory interactivity: active learning approaches, for example, have been found to improve recall [27–29]. Further work is needed to fully characterize the space of trade-offs in implementation complexity and benefit to reader understanding across different types of narrative and interactivity in communicative visualization. Stepping back, we question whether exploratory interactivity should be considered an essential component of communicative visualizations, rather than a component to be employed as and when needed.

2 Background and Motivation

A number of slightly differing definitions of narrative and interactivity in visualization have been put forth in the literature. We look at these different definitions and identify easy-to-operationalize definitions that we can use to systematically design new visualizations.

2.1 Narratives in visualization

According to Bach et al. [7], there is an important distinction between the terms *story* and *narrative*, which are often used interchangeably in the context of communicative visualizations. A story is the set of facts or data which is presented directly to or inferred indirectly by the reader. On the other hand, a narrative is an author-defined sequence of events with the goal of making facts or data clear and compelling to the reader. In other words, a particular narrative is one of many ways of presenting a given story to the reader.

Segel and Heer first defined a design space for visualizations which employ a narrative, and placed them along a spectrum between reader-driven and author-driven approaches [44]. According to them, *author-driven* visualizations have little or no interactivity and rely heavily on messaging and ordering, to form the narrative; on the other hand, *reader-driven* visualizations are highly interactive and have very little or no messaging and ordering. They identified three commonly used schemas on this spectrum: *martini glass structure* (primarily author-driven), *interactive slideshow*, and *drill-down story* (primarily reader-driven).

Subsequently, there have been several attempts at more precisely defining narrative visualizations. Kosara and Mackinlay defined a narrative as *an ordered sequence of steps with a clearly defined path through it*. When the steps primarily consist of information visualizations, they are termed as narrative visualizations [31]. Hullman and Diakopoulos define narrative visualization as a style *which combine persuasive, rhetorical techniques for explanation with interaction techniques for exploration by the user* [24].

Lee et al. [32] state that a *visual data story* should (1) consist of a set of facts backed up by data; (2) be visualized to support one or more intended message, and can include annotation or narration to clearly highlight and emphasize this message and avoid ambiguity; and (3) be presented with a meaningful order or connection between its parts to support the author's high level communication goal. Although Lee et al. use the term *visual data story*, it is clear that they are referring to the genre of narrative visualization.

Bach et al. [7] identified several *narrative design patterns* commonly used in communicative visualizations to create a compelling narrative. They define narrative design patterns as: "low-level narrative devices that serves a specific intent. It can be used individually or in combination with others to give form to a story". These high-level strategies can be (but are not limited to): (1) the use of long-form text, annotations, labels etc. [32]; and (2) ordering that ties together the facts presented to the reader [25].

In our study, we examine the effect of an author-defined narrative with messaging and prescribed ordering, which we term as an *explanatory narrative*. We designed versions of the same visualization to clearly have or not have explanatory narrative. A visualization with explanatory narrative uses short pieces of text to present insights from the data to the reader, and uses an author-specified sequence to present these pieces of text. This does not represent all approaches to narrative visualization, but is one definition that is straightforward to systematically vary. We are interested in the effect of this overall explanatory narrative strategy (not specific narrative design patterns in particular); thus, in our designs, a visualization which does not have a explanatory narrative does not have any form author-specified ordering, messaging or interpretation of the facts / data. We include these assumptions in our analysis to account for variance that may arise due to different instantiations of the overall explanatory narrative approach (see Section 4.6).

2.2 Interactivity in visualization

Although interactivity is commonly used in information visualization, there have been continued efforts to define precisely what constitutes *interactivity*. Some of these definitions are derived from how interaction techniques are defined for human–computer interaction (HCI): "The interaction component involves the dialog between the user and the system as the user explores the dataset to uncover insights. The interaction component's roots lie in the area of HCI." [50].

Interactivity is often used to allow readers to explore, dig deeper into, discover insights from, and make sense of data [34, 50]. According to Yi et al., this can be categorized into the following seven types of interactions: *select, explore, reconfigure, encode, abstract/elaborate, filter*, and *connect*. On the other hand, Ziemkiewicz and Kosara make a distinction between trivial and non-trivial interaction, where non-trivial interaction facilitates *active reading* (allowing the reader to actively seek information) by allowing them to make changes to the parameters of the visual mapping [52]. We use this distinction to inform what constitutes *exploration* through interactivity in visualization: exploratory interactivity should only include operations the user performs on a visualization to seek information, rather than merely accessing information presented to them by the designer.

Alternatively, we consider interactions which do not change the visual mapping in any way and do not allow the user to explore the

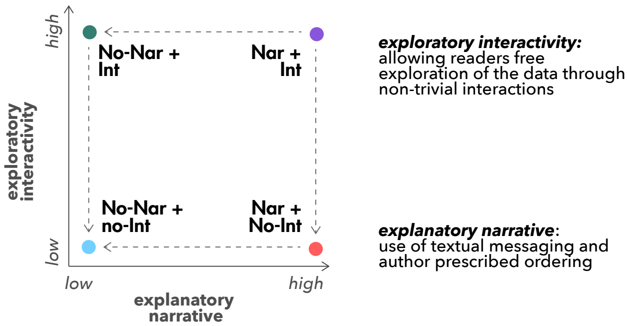

Fig. 1. Design space for communicative visualization with two complementary axes of explanatory narrative and exploratory interactivity. This space is a subset of a larger space of explanation and exploration proposed by Thudt et al. in [46].

data beyond what has already been presented to them to be *trivial*. For example, some hover interactions, which fall under the category of *elaborate*, are used to give more precise information about the data already depicted in the visualization without showing additional dimensions in the dataset. Interactivity has also been used to incorporate active learning techniques into visualizations through tasks such as prediction [4,27–29]. Kim et al. have found such approaches to improve users' recall of the data. However, in our study, we do not consider these types of interactivity as they do not enable user exploration.

In our study, we focus on visualizations which allow the reader to *explore* the data through *non-trivial* interactions, and term this class of visualizations as *exploratory interactive* visualizations. Similar to narrative design patterns, we do not control for the use of different categories of interactions, and instead focus on the broader presence or absence of exploratory interactivity. We assume that there is variance in the application and effects of these specific techniques, which we build into our model as random effects (Section 4.6).

2.3 Design space for communicative visualizations

Thudt et al. in [46] propose a design space for communicative visualizations that consists of two (potentially complementary) dimensions: explanation and exploration. We adapt this design space to consider two design strategies—*explanatory narrative* and *exploratory interactivity*—using the definitions of each strategy established above (Fig. 1). Our adaptation is a constrained version of the design space proposed by Thudt et al., as we only consider specific forms of explanation (explanatory narrative) and exploration (exploratory interactivity). One could imagine, for example, explanatory interactivity, which might take the form of active learning approaches [4, 27–29]; we do not consider such designs here.

Since our goal is to study the effect of these two visualization design strategies, we vary the presence/absence of *explanatory narrative* and *exploratory interactivity* to create our four conditions: *no-Nar+no-Int* (baseline condition), *Nar+no-Int*, *no-Nar+Int*, and *Nar+Int*; which roughly correspond to the points A, B, C, and D in Fig. 1. We describe this design process in Section 3.

2.4 Evaluation of communicative visualizations

The goals of communicative visualizations, which target a broad group of users and have varying usage patterns [39], can be many: memorability, enjoyment, engagement, transfer of knowledge, etc. [36, 41]. There has been a growing body of work in information visualization that goes beyond testing cognitive efficiency, accuracy, and other usability goals [41] to evaluate memorability [14, 15], engagement, [8, 16, 22] and the ability of the user to store the data presented in memory [8, 13, 22].

Bateman et al. found that the use of *chart junk* (or embellishments) in visualizations can help participants perform better on long-term recall tasks involving factual knowledge [8]. Borkin et al. found that features such as human-recognizable objects and the use of more colors can make a visualization more memorable and allow people to recall more details about the visualization [14]. Haroz et al. found that participants were able to recall data just as accurately in charts with pictographs as in charts with simple shapes (such as bar charts) [22]. However, these studies either used a descriptive *free-recall* [8, 14], where participants were asked to describe as many details about the chart as possible, or recall of the data values from the chart [22].

Borgo et al. [13] found some evidence that visual embellishments in static charts "can help participants better remember the information presented in the visualization" in terms of working memory, long-term memory and *concept grasping*, which involves identifying key concepts depicted in the visualization. Borgo et al. used multiple-choice questions to measure participants' ability to learn facts and concepts presented in the visualization. Based on their study design and types of visualization used, each participant had to identify the primary concept or piece of information that was communicated through the graph.

The studies above elicited free-recall through open-ended questions [8, 14, 22] or used multiple choice questions [13] to assess concept grasping. We adopt a *learning assessment* approach from the education literature: the revised Bloom's taxonomy [6, 12], which provides a mechanism to systematically assess learning. It defines two dimensions: the knowledge dimension and the cognitive process dimension. The knowledge dimension constitutes of different types of knowledge such as *factual* and *conceptual*. The cognitive process dimension consists of different sets of cognitive processes which can be used to assess learning at a particular knowledge dimension through a corresponding set of tasks such as *recall*, *recognition*, *classification* etc.

In our study, we use the revised taxonomy to devise a set of true / false questions to assess recall of factual knowledge from each visualization. This allows us to measure the effect of explanatory narrative and exploratory interactivity in a communicative visualization on learning.

The effect of a narrative component (though not necessarily an explanatory narrative as we define it) on user behavior has been previously investigated by Boy et al. [16]. In their study, participants were presented with an exploratory interactive visualization, with or without the presence of an introductory narrative preceding the visualization. They measured time spent on the visualization and the number of "meaningful" interactions (hover and click) with it—two behavioral metrics commonly used to measure engagement [20]. Based on these metrics, they found that the presence of an introductory narrative component led to slightly more time spent on, but not in the number of interactions performed with the visualization. Hence, they concluded that the narrative component may not help increase understanding of the data.

However, an implicit assumption in Boy et al.'s study [16] is that more exploration will lead to more understanding. While these factors may be related, there are good reasons to believe that engagement is at best an imperfect proxy for learning: just because a user spends time digging through data does not mean they will find useful knowledge; in the context of communicative visualization (where a particular story might be important), it is even less likely that through exploration a reader will find the (or any) story of interest. There is even potential to backfire: for example, researchers have found users of exploratory analytic systems to be susceptible to drawing spurious conclusions [51].

Simple engagement metrics are easy to record, but may not adequately describe user behavior [20], and we lack any studies on the validity or reliability of engagement metrics as a proxy for learning metrics like recall. Through two exploratory analyses, we attempt to get an initial understanding of how well metrics for measuring user behavior can be used as proxies for learning. First, we evaluate if the presence or absence of explanatory narrative and exploratory interactivity has an effect on each user behavior metric. This is a replication of the original study to see if our participants behave in a similar way. Second, we calculate if metrics of user behavior are correlated with answering questions correctly, to see if the tendency to explore leads to greater recall of the data.

3 Design of Visualizations

Using our operational definitions of explanatory narrative and exploratory interactivity, we developed a design process to systematically create different variants of the same visualization that have or do not have explanatory narrative or exploratory interactivity. We applied that process to existing, professionally-produced visualizations in to create stimuli for our experiment.

3.1 Selecting candidate visualizations

We first created a catalog of professionally-produced visualizations by going through online visualization galleries, such as Visualizing.org[2], and websites of news agencies that maintain lists of visualizations published by their graphics departments, such as The New York Times, The Washington Post, and Bloomberg. We restricted ourselves to professionally-produced visualizations to ensure high design quality. We identified 40 visualizations, and identified the position of each in our design space (Fig. 1). From this catalog, we selected four visualizations [19, 23, 37, 40] using the following criteria: (1) they possessed a strong narrative component; (2) they were suitable for adaptation to the other corners of the design space; (3) they depicted data which was readily and publicly available for reproduction. We selected four visualizations to increase the chance that any observed effects would be consistent across multiple distinct visualization designs and not be specific to the properties of one particular visualization or dataset.

3.2 Visualization re-design process

Each corner of our design space (Fig. 1) represents one condition we want to evaluate: *no-Nar+no-Int*, *Nar+no-Int*, *no-Nar+Int*, and *Nar+Int*. Therefore, for each of our four visualizations, we created four design variants, one for each corner of our design space. This resulted in $4 \times 4 = 16$ unique combinations of visualization type and condition.

We used a systematic design process to adapt each visualization to each corner of the design space. We first took the visualization's original version and determined its position in the design space. We deconstructed the existing components of the visualization (e.g., narrative and interactivity components), and created paper-based mock-ups of the visualization. Then, through an iterative design process, we took systematic steps along one axis or the other (explanatory narrative or exploratory interactivity) to create variants of the same design. Fig. 2 shows this high-level redesign process, and Fig. 3 shows an example of the redesign of one of our visualizations, the *Gun deaths* visualization (each visualization is described in more detail below). We designed each version to be as consistent as possible to the original, with consistent encoding and layout across each version of a visualization type.

The *Nar+Int* version of each visualizations used a Martini-glass structure [44]—consisting of a narrative section, which used either scrolling or steppers to advance the narrative, followed by an exploratory interactive section. The *Nar+no-Int* version consisted of just the narrative section of the corresponding *Nar+Int* version. Similarly, the *no-Nar+Int* version consisted of the exploratory interactive section of the corresponding *Nar+Int* version. Finally, the *no-Nar+no-Int* version had neither interactivity, messaging, nor prescribed ordering to ensure the absence of explanatory narrative.

We implemented all 16 versions as standalone web-pages using D3.js. Since we use recall of factual knowledge as our assessment metric, we needed to ensure equal expressiveness: the visual encoding of the data should express an equal amount of information for each variant of a visualization [38]. To verify that each variant had the same degree of expressiveness, we conducted an *informal pre-test pilot study* with nine graduate students of HCI of a large public university. We presented each student with four visualizations—one version for each visualization type and condition—and asked them to answer the assessment questions (a true/false questionnaire; see Section 4.3). We made iterative changes to the design based on their feedback. We stopped when two successive participants were able to answer all the assessment questions correctly for each version.

Since we cannot test all visualization designs, to decrease the likelihood that any findings are specific to one particular type of visualization, we test four distinct visualizations. This helps us identify the effect of explanatory narrative across different visualization designs and identify if our findings are generalizable. However, this does not imply that we assume the effect of explanatory narrative and exploratory interactivity to be the same for each visualization type. Rather, by using varying slopes and intercepts in a multilevel model, we will estimate different effects for the presence or absence of explanatory narrative and exploratory interactivity across the different types of visualizations, and the variance of those effects (see Section 4.6).

3.3 The visualizations

In what follows we will briefly describe the 4 visualizations. The visualizations have been included in the supplementary materials and require a local server to be setup. They can also be accessed here: `https://visrecall.github.io/`. The contents of the website have been anonymized for review and we do not track IPs or any visitor information.

Gun deaths [19] explores the 30,000+ annual deaths in the United States which are caused by gun violence using a dot matrix plot, where each dot represents a single victim. Other information encoded includes the type of death and demographics. The original graphic maps onto the *Nar+Int* corner of our design space, using steppers to allow the reader to advance the narrative; the final step is the interactive section, users can explore the data using a set of filters. We create the *Nar+no-Int* and *no-Nar+Int* versions by keeping only the respective explanatory narrative and exploratory interactivity sections from the *Nar+Int* version. The *no-Nar+no-Int* version is created by showing several graphs broken down by type of death and demographic.

Carbon clock [40] is a time-series line chart visualization, depicting atmospheric CO_2 trends for the past 3 years, 60 years, 12,000 years and 800,000 years. It maps onto the *Nar+Int* corner of the design space. The narrative section uses "scrollytelling" (continuous scrolling where the visualization is positionnaly fixed on the screen and the text updates as the user scrolls); the interactive section allows users to toggle through the views (past 3 years, 60 years, 12,000 years or 800,000 years) using buttons. We create the *Nar+no-Int* and *no-Nar+Int* versions by keeping only the respective explanatory narrative and exploratory interactivity sections from the *Nar+Int* version. The *no-Nar+no-Int* version consists of a separate graph for each time period.

Measles [23] is an animated visualization depicting how ten hypothetical communities with different vaccination rates will be affected when they come into contact with a person infected with measles. The original visualization maps on to the *no-Nar+no-Int* corner and is accompanied by a passage describing the graphic, the assumptions made, and the 2015 outbreak of measles in the US. We use this passage to create an explanatory narrative for the *Nar+no-Int* and *Nar+no-Int* conditions. We use the county-level US kindergarten MMR vaccination rate assessment data [30] to create the exploratory interactive section (*Nar+Int* and *no-Nar+Int*) by enabling users to change the vaccination rate using a slider and re-run the simulation.

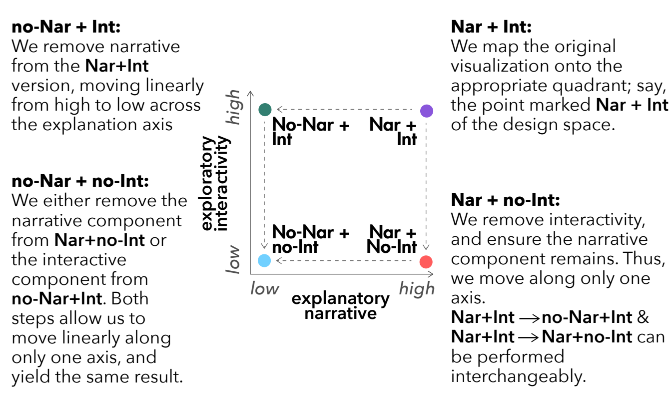

Fig. 2. We use our design space (Fig. 1) to systematically adapt the original visualizations to each of our four conditions. How this adaptation proceeds depends on where in the design space the original version was; see Fig. 3 for an example of the adapation process.

[2]https://www.visualizing.org/

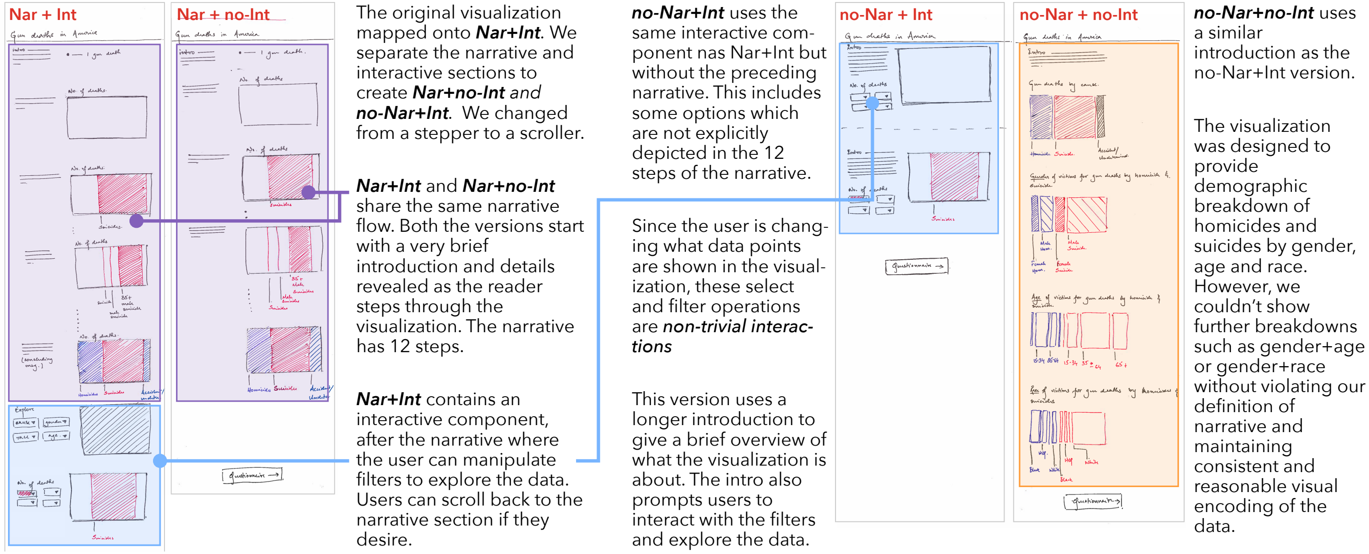

Fig. 3. Adapting a visualization from its original place in the design space to other corners of the design space. Here, we show the adaptation process for *Gun deaths* visualization. The original version mapped onto Nar+Int corner. We iteratively designed versions which mapped on to other corners using the definitions of narrative and interactivity. We developed using HTML, JavaScript and D3.js

Healthcare [37] is a two axis parallel coordinate visualization which compares the cost of healthcare along 7 different metrics to the corresponding quality of care measured by 8 metrics in 35 OECD countries. It mapped onto the *no-Nar+Int* corner of our design space. Users can use two drop-down menus to change the metrics displayed along the two axes. An introductory passage precedes the visualization and introduces the reader to the data, which we use to create a explanatory narrative for the *Nar+Int* and *Nar+no-Int* versions.

4 Method

4.1 Primary research questions

Through this study we attempted to answer the following, pre-registered[3], primary research questions:

1 (a) What is the effect of *explanatory narrative* on recall?

1 (b) What is the effect of *exploratory interactivity* on recall?

For both the research questions, we use the probability of an average participant answering an average question correctly is as the dependent measure. The primary predictors are *interactivity* (present or absent) and *narrative* (present or absent), along with their interaction term.

4.2 Exploratory research questions

Two subsequent questions arise from our primary research question, prompting us to conduct exploratory (non-preregistered) analyses. First, we conduct a replication of Boy et al.'s study to evaluate if the presence of explanatory narrative or exploratory interactivity affect readers' exploratory behavior. We investigate:

2 (a) What is the effect of the presence of *explanatory narrative* on the total time spent on the visualization?

2 (b) What is the effect of the presence of *exploratory interactivity* on the total time spent on the visualization?

Second, even if readers engage with an exploratory interactive component of a visualization there is no guarantee that this leads to greater learning. Thus, we evaluate whether engagement metrics [16] are correlated with learning measured by the number of correctly answered questions by each participant:

[3]The anonymous pre-registration document can be found in the supplementary materials

3 (a) What is the effect of total time spent on the visualization on recall?

3 (b) What is the effect of the number of non-trivial interactions on recall?

4.3 Assessing learning

To answer our primary research questions we created a ten-item true / false questionnaire for each visualization[4]. We used the revised Bloom's taxonomy [6], which allows us to focus on assessing the type of cognitive process (in our case, recall) rather than the subject area (topic of the visualization) [1]. The assessment tasks associated with recall include recall and recognition of facts, data, and basic ideas or concepts presented through the visualization.

To develop the assessment questions for a given visualization, we identified ten pieces of factual knowledge presented through that visualization. We then created one true/false question for each of these ten facts. We ensured that all the questions could be answered by participants in any condition through our pilot study (Section 3.2). We created one additional question to each questionnaire as an attention check, to which the answer was clearly false if the reader understood the topic of the visualization. For example, the attention check question for the *Gun deaths* visualization was: "None of the deaths shown in the graphic were caused by an incident which involved a gun".

4.4 Other metrics

To answer our exploratory research questions, 2(a), 2(b), 3(a), and 3(b), we also collected data on the amount of time spent by participants on the visualization webpage. In the two conditions with exploratory interactivity (*no-Nar+Int* design, *Nar+Int*), we also measured the number of non-trivial interactions with the data: the number of clicks used to filter the data (Gun deaths), change the view (Carbon clock), change the vaccination rates using an input slider (Measles), or change the factors being visualized (Health care).

4.5 Study design and procedure

We used a between-subjects design where we randomly assigned each participant to a single condition—each participant was presented with one visualization and then answered the corresponding assessment questionnaire. Participants were not allowed to go back to the visualization once they reached the questionnaire, and were informed of this at

[4]The questionnaires can be found in the supplementary materials

the beginning of the study. We do not collect any demographic information, and do not attempt to balance covariates, relying on randomization and modelling for proper inference. According to Althouse et al. [5], baseline balance in all covariates is not necessary for valid statistical inference: "under proper random treatment assignment, distributions of all baseline covariates among treatment groups are random...As Senn [18-21 in 5] has discussed previously, the standard probability calculations applied to [randomized] clinical trial results already make an allowance for the fact that the treatment groups will almost certainly be imbalanced.".

We performed a prospective power analysis to determine the number of participants to recruit. Based on this, we decided to recruit 100 participants per condition. We pre-registered our Bayesian regression model, along with exclusion criteria using AsPredicted.org before collecting and analyzing our final dataset. We launched the study as a single HIT on Amazon's Mechanical Turk (MTurk). Participants were instructed to go through each step of the visualization carefully and then proceed to the questionnaire.

On MTurk, we recruited participants who have a prior HIT approval rating of 98% and have completed at least 500 prior HITs. Each participant was given a base pay of $0.75 and informed that they would receive a bonus of $0.2 for every question that they answered correctly. We introduced the incentive for answering questions correctly to motivate participants to spend time on the visualization, with the goal to simulate the intrinsic motivation that an internet user might have to read a visualization on a news site. The average time to finish the HIT was slightly under 8 minutes and the average payoff was $2.30. In total we received 389 responses: 97 (83)[5]in static, non-narrative condition; 100 (86)[5] in interactive only condition; 105 (95)[5] in the narrative only condition; 87 (80)[5] in interactive and narrative condition.

Our pre-registered exclusion criteria were: (1) answering *True* to all or *False* to all questions, and (2) failing to answer the attention check question correctly. We rejected two participants outright as they answered *True* to all questions. Another 43 participants failed the attention check question. We noticed that a majority of the participants (31) who failed just the attention check question saw the Carbon clock visualization, which was: *the graphic showed the amount of Carbon Dioxide in Earth's water bodies - oceans, seas and rivers - has been increasing*. We suspect that this attention check question may have been more difficult than the others. Hence, we performed our primary analysis twice—first excluding the participants according to our pre-registration, and then including the 43 participants who were excluded for failing any attention check question. The results for both the analyses were similar, and therefore we report our pre-registered analysis in this paper. The results of both analyses can be found in the supplementary materials.

4.6 Model

After Kay et al. [26], we implemented a Bayesian multilevel logistic regression model for our primary analysis using the *brms* package in R [17]. Our model can be represented using the *lmer* formula syntax [9] as follows (refer to the supplementary materials for the complete model specification):

$$
\begin{aligned}
logit(correct) \sim\ & interactivity \times narrative \\
& + (1|participant) + (1|question) \\
& + (interactivity \times narrative|visualization)
\end{aligned}
$$

Our model estimates the probability of an average participant answering an average question correctly, using *explanatory narrative* and *exploratory interactivity* and their interaction as population-level (fixed) effects. We include group-level effects for *participants*, *questions* and *visualization*.

We randomly assign participants to each study condition. However, we cannot assume that all participants would have equal baseline knowledge. Instead, it is more likely that participants will have differing abilities resulting in different baseline probabilities of answering a question correctly. To account for these possible differences, we model each *participant* using a varying intercept.

We use different questions in our assessment questionnaire, and all questions may not be equal in difficulty, which may result in the difficulty of the set of questions for a visualization to vary. We model each question using a varying intercept, where each *question* has an unique intercept based on its difficulty.

Finally, the effect of *narrative*, *interactivity* and the interaction term may also be different for each *visualization*, because the explanatory narrative and exploratory interactivity used in each visualization are not identical. This may be due to the presence of different narrative design patterns or interaction categories. However, the explanatory narrative and exploratory interactivity techniques for a visualization also shares some common aspects with the explanatory narrative and exploratory interactivity techniques for other visualizations, which allows us to broadly classify them as explanatory narrative and exploratory interactivity. When trying to estimate the effect of *explanatory narrative*, *exploratory interactivity*, this is an important consideration which we take into account through our hierarchical model, by using varying slopes and intercepts for these two variables, for each visualization.

5 Results

5.1 Primary analysis

We show the posterior probability density, posterior mean, and 66% and 95% quantile credible intervals for the probability of answering a question correctly. In Fig. 4.1, we show the estimates, averaged over the questions we have tested in our questionnaire. This estimate averages the variance for different questions, and therefore the estimates are more precise and close to the observed proportions.

In Fig. 4.2, we show the estimates for a typical (or "average") question, which take into account the variance in questions that we have seen and the difference in difficulty of different questions. As a result, these estimates will be more uncertain, and may not correspond

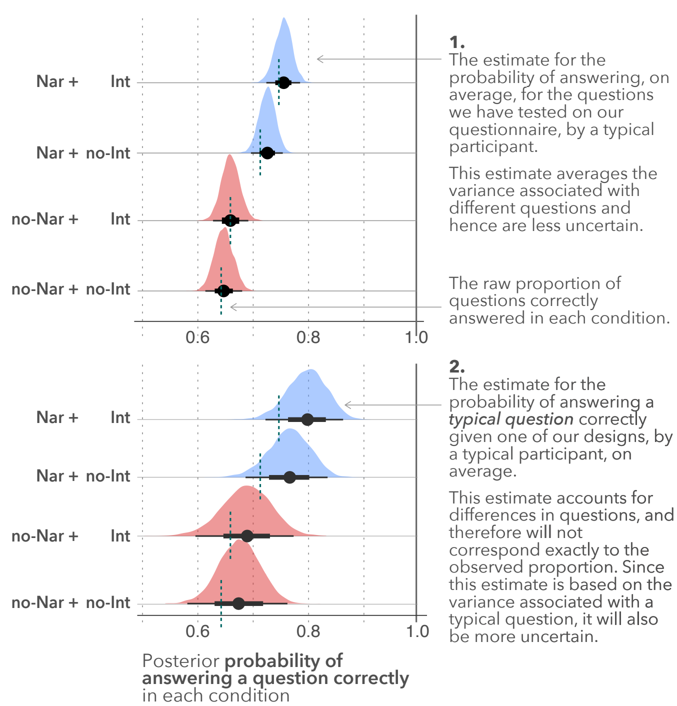

Fig. 4. Posterior probability density, posterior mean, 66% and 95% credible intervals of p (the probability of answering a question correctly).Credible intervals are the Bayesian analog of confidence intervals, and indicate where, based on the data and prior distributions, the posterior estimate of p may lie. We marginalise (average) over the group-level effects of visualization. The results shown excludes participants who failed the attention check (preregistered criteria).

[5]numbers in parentheses indicate the number of responses after removing participants based on our pre-registered exclusion criteria.

Fig. 5. Posterior probability densities for the mean differences between the conditions. We also show the group-level effects for each visualization, as well as the expected effect for a new visualization.

exactly to the observed proportions (as the observed proportions come from questionnaires with multiple questions having varying difficulty), though the differences in proportions should be similar. Acknowledging the uncertainty present, it is more important to discuss our results for a typical question, and not specifically to the questions that were used in our questionnaire. Thus, we will refer to the estimates for a typical question subsequently.

Fig. 5 shows the mean difference for the effect of *exploratory interactivity* and *explanatory narrative* on the probability of answering a typical question correctly. We find that the presence of a explanatory narrative has a small but positive effect on recall—explanatory narrative increases the probability of answering a question correctly on average by 10 percentage points (95% CI: [6.1, 14.3]) (top portion of Fig. 5). The mean effect size is of the order of getting one more question correct on the ten item questionnaire, for an average participant, for a particular visualization. This effect is fairly consistent across visualization type (middle portion of Fig. 5). On the other hand, the presence of exploratory interactivity likely has little or no practical effect—exploratory interactivity increased the probability of correctly answering a question by 2.3 percentage points (95% CI: [–1.4, 6]). The upper end of this interval represents only about half a question more correct on average, and the effect is most likely smaller than that.

As described in Section 4.6, the effect of explanatory narrative and exploratory interactivity vary for each visualization, and we account for this using varying slopes and intercepts. Even given that variance, the presence of explanatory narrative has a fairly consistent positive effect across visualizations. The effect of exploratory interactivity is also fairly consistent and is unlikely to be at or greater than 10 percentage points in any of the visualizations. Our model also allows us to generate *predictions* for the mean effect on a new visualization by incorporating the variance of the visualization-level effects into the estimate of the mean difference (bottom part of Fig. 5). These distributions suggest that given a new visualization we have never seen before, there is some chance that explanatory narrative and exploratory interactivity could have more or less impact *for that particular visualization* than what we observe here.

5.2 Exploratory analysis

To answer our exploratory questions 2 (a) and 2 (b), we fit two multilevel models with the total time spent on the visualization and the

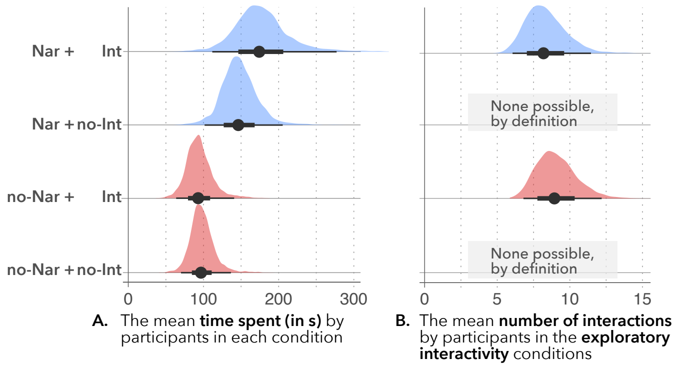

Fig. 6. Posterior probability densities for the time spent, and the number of interactions performed with by an average participant in each condition

number of interactions performed with the visualization as the dependent variables respectively; *explanatory narrative* and *exploratory interactivity* were the independent variables for both models, with varying slopes for each visualization. Fig. 6 depicts the results, which shows that the presence of explanatory narrative may have a small positive effect on time spent; the presence of exploratory interactivity likely does not have a positive effect on time spent. We also observe that the presence of explanatory narrative likely does not result in users performing more non-trivial interactions with the visualization. However, we should note here that there is a lot of uncertainty in our estimates, as evidenced by the wide 95% intervals.

To answer our exploratory questions, 3 (a) and 3 (b), we fit two multilevel linear models to estimate the effect of time spent and number of interactions performed on the number of questions correctly answered. We find a weak log-linear relationship between the duration of time spent on the visualization as the independent variable for the first model (Fig. 7A). 7A indicates that there may be a slight positive correlation of time spent with the number of questions correctly answered by our participants. However, in Fig. 7B we see that there may not be any positive correlation between the number of non-trivial interactions and the number of correctly answered questions.

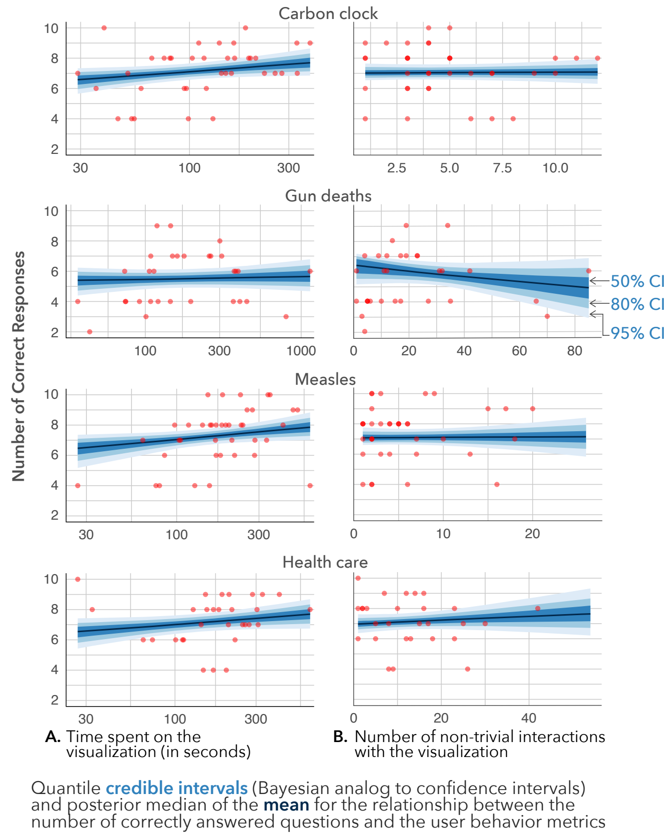

Fig. 7. The effect of (A.) the total time spent in the visualization and the (B.) the number of interactions performed by a participant on the number of questions answered correctly

6 Discussion

6.1 Effect of communicative visualizations on learning

Communicative visualizations can have many goals. Our work focuses on the effectiveness of such visualizations in helping readers learn knowledge that the author intends to convey. Our findings indicate that the presence of explanatory narrative can have a small positive effect on learning, as evidenced by the increase in the probability of answering recall questions correctly (Fig. 5B). Moreover, the size and direction for this effect is fairly consistent for each visualization type we used (middle portion of Fig. 5B). We attempt to acknowledge the uncertainty in how this effect might translate to new visualizations we have not seen before through our multilevel model: certainly, some uses of explanatory narrative will be ineffective (or even detrimental to understanding), and some may be extraordinarily good, but we expect typical usage to yield a modest, positive effect.

One explanation for our findings may be that the use of explanatory narrative helps the designer convey information in a concise and crystallised form, which helps readers focus on the most important aspects of the data presented in the visualization. Thus, when disseminating information to a large audience, which may comprise of diverse user groups and involve different usage patterns, we believe that designers adopting explanatory narrative techniques can make learning more efficient and effective.

On the other hand, the presence of exploratory interactivity in a communicative visualization likely does not have a large positive effect on learning (Fig. 5A). Hypothetically, exploratory interactivity should allow readers to freely explore the data presented and discover meaningful relationships. However, simply providing exploratory interactivity does not mean readers will take advantage of it, and even if readers do take advantage of it, they may not (through happenstance or variance in ability) come across any meaningful relationships in the data.

Nevertheless, given the small (though uncertain) positive effect of exploratory interactivity, it may be that some readers who perform more interactions are able to learn more information. In our exploratory analysis, we looked at the effect of *the duration of time spent* and *the number of interactions performed* on recall, which revealed mixed results. We found a weak positive correlation between time spent on the visualization and recall. However, we found that there was likely little or no correlation between the number of interactions performed in the exploratory interactivity visualizations (*no-Nar+Int* and *Nar+Int* conditions) and recall.

Thus, for communicative visualizations, we find some evidence to suggest that enabling un-directed, interactive exploration of the visualization may not be the most effective way to help readers learn. This may be because such exploration may not result in the reader uncovering all the relevant information from the visualization (particularly when graphical literacy varies), or even worse, may result in misinterpretations or spurious conclusions [51].

6.2 Costs and benefits in designing communicative visualizations

To inform the design of communicative visualizations, it is worth considering the costs and benefits to employing interactivity and narrative. Creating interactive visualizations which function seamlessly across multiple types of devices and interaction modalities can be challenging and expensive for an organization—design and implementation requires more time than for a static visualization, and development may even require hiring additional personnel with expertise in interactive visualization. Similarly, creating a narrative can entail costs as well. As not all uses of narrative will be effective, creating a compelling narrative might require expert storytellers; Extracting meaning from the data may require expert analysts. However, for news organisations, we expect that such expertise may be more readily available.

In light of the relatively small potential benefits to exploratory interactivity we have found, if a designer's goal is to communicate a particular message to their readers, the potential benefits to exploratory interactivity may not be worth the associated design, implementation, and personnel costs. This cost/benefit ratio may help explain explain why fewer outlets have been creating purely exploratory interactive news graphics.

However, exploratory interactivity does not represent the only type of interactivity used in communicative visualizations. Kim et al. have shown that using interactions to elicit users' prior knowledge and providing feedback can improve data recall and comprehension [27–29]. These approaches use interactivity to support *active learning*, and have been employed in data journalism; e.g. *The New York Times*' "You Draw It" visualization [4]. For communicative visualization, perhaps emphasis should be placed on interactivity specifically designed to improve learning from a visualization, instead of augmenting visualizations to allow freedom of exploration, as in the exploratory interactivity approach we tested. This suggests that the purposes and techniques for interactivity in communicative visualization may be quite different from those traditionally developed for information visualization for exploratory data analysis.

Finally, even though a large percentage of the consumers of communicative visualizations do not use exploratory interactive features when they are present [2,47], and our analysis suggests that exploratory interactivity does not substantially improve learning, this does not necessarily imply that it has no benefit. Exploratory interactivity may help provide transparency and increase trust in the data and the source—Aisch posits that since all of the data is accessible to the audience, skeptical viewers may be more certain that a visualization is not depicting partial results or hiding important aspects of the data [3]. These claims about the potential benefits to adding interactivity in such contexts have, to the best of our knowledge, not been empirically investigated, and may be a possible avenue for future research.

6.3 Using engagement metrics to evaluate visualizations

In our exploratory analyses, we attempted a replication of Boy et al.'s [16] study for communicative visualization, and our results are similar. Note that however, unlike their users, who were either *information-savvy* or *visualizaiton-savvy*, our users were Mechanical Turkers.

It is worth addressing discrepencies in the conclusion between our work and Boy et al. [16]. Both our and their paper had similar results with respect to attention span, finding that given readers' limited attention spans, when a narrative section is provided and people spend time on it, they are consequently less likely to spend time on an exploratory section. However, the stated design goal in Boy et al. was to encourage use of the exploratory section of their visualizations: "our goal is to understand how to engage people with exploratory visualizations on the web"; perhaps on the assumption that making use of an exploratory section is likely to lead to more knowledge discovery. Therefore, given that adding a narrative section led to less use of the exploratory section, they conclude that one **should not** employ an narrative techniques.

By contrast, our goal is explicitly knowledge transfer in the context of a communicative visualization (which we measure directly), not engagement or exploration (which we consider useful only insofar as they serve the goal of knowledge transfer). We similarly find that adding a narrative section is likely to steal some attention from the exploratory section, but also that an explanatory narrative likely improves knowledge transfer. Therefore, since we do not consider use of the exploratory section an end in itself (but rather a technique to be employed if it improves knowledge transfer), we conclude that one **should** employ narrative techniques.

Further, in our second exploratory analysis, we observe that the number of correctly answered questions has a weak positive correlation with time spent (Fig. 7A), but little or no correlation with the number of interactions performed. This suggests that simple metrics of user engagement may not be adequate proxies for other goals, such as learning.

Mahyar et al. [35] note that "there is no clear definition of what engagement means in the Information Visualization domain". Recent work has proposed alternative metrics for measuring user behavior such as *exploration uniqueness*, *exploration pacing* [20], and *data point coverage*, *data point distribution* [48] etc. Although these proposed metrics may better describe users' exploratory interactions with visualizations, further research is necessary to identify what constructs these behavioral measures best correspond to: are these measures of engagement, learning, or something else?

Until such validation has been done, we recommend that researchers making use of engagement metrics be very clear about their research goals: engagement itself may be central to one's research questions, in which case engagement metrics are an obvious fit. However, higher engagement according to these metrics does not necessarily imply more learning; if one's research questions relate to learning, it may be better to measure learning more directly.

6.4 Limitations

The questions we tested participants on naturally do not encompass all the information that can be gathered from each visualization. Our study is concerned with the message that the visualization is trying to convey, and we focus only on recall, which is just one cognitive process involved with learning. Our study design does not allow us to investigate other phenomena associated with learning from visualizations: for example, some readers may have come up with their own questions and discovered answers to these questions. Additionally, we do not measure higher-order cognitive processes associated with learning described in Bloom's taxonomy [6, 12], such as *apply*, *analyse*, *evaluate* and *create*. Instead, one approach might be to conduct a qualitative study, similar to the methods proposed by Saraiya et al. [42, 43]. This may be more effective in comparing *all* the information that the reader of a visualization gleans from it, rather than the specific message intended by the designer.

7 Conclusion

We study the effect of explanatory narrative and exploratory interactivity in communicative visualization on learning, measured using recall of the data presented in the visualization. We find a consistent small but positive effect on recall due to explanatory narrative, but likely little meaningful effect due to exploratory interactivity. We also observe that time spent on the visualization may have a weak positive correlation with recall, but the number of interactions with the visualization (a metric which is commonly used to measure engagement) may not be correlated with recall. Thus, we recommend caution in using engagement metrics as a proxy for learning; if learning is central to one's research questions, better to measure it more directly.

In light of our results, we argue that exploratory interactivity should not be considered an essential component of communicative visualizations. Instead, if the main goal of a visualization is to convey a particular message to the viewer, emphasis should be placed on effective narrative and visualization techniques (including interactive visualization techniques) designed to support learning. Put another way, if communicative visualization design is viewed as a user-centered problem, interactivity should be added only if it helps the design support users' learning, not assumed *a priori* as a necessary component.

References

[1] P. W. Airasian and H. Miranda. The role of assessment in the revised taxonomy. *Theory into practice*, 41(4):249–254, 2002.

[2] G. Aisch. Gregor Aisch, Information+ 2016 Conference presentation. Video, September 2016.

[3] G. Aisch. In Defense of Interactive Graphics. vis4.net, March 2017.

[4] G. Aisch, A. Cox, and K. Quealy. *You Draw It: How Family Income Predicts Children's College Chances*, 2015 (accessed March 28, 2019).

[5] A. D. Althouse, K. Z. Abebe, G. S. Collins, and F. E. H. Jr. Response to "Why all randomized controlled trials produce biased results". *Annals of Medicine*, 50(7):545–548, 2018. PMID: 30122065. doi: 10.1080/07853890.2018.1514529

[6] L. W. Anderson, D. R. Krathwohl, P. Airasian, K. Cruikshank, R. Mayer, P. Pintrich, J. Raths, and M. Wittrock. *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy*. Longman Publishing, 2001.

[7] B. Bach, M. Stefaner, J. Boy, S. Drucker, L. Bartram, J. Wood, P. Ciuccarelli, Y. Engelhardt, U. Köppen, and B. Tversky. Narrative Design Patterns for Data-Driven Storytelling. In N. H. Riche, C. Hurter, N. Diakopoulos, and S. Carpendale, eds., *Data-Driven Sotrytelling*, pp. 107–133. CRC Press, 2018.

[8] 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*, pp. 2573–2582. ACM, ACM, 2010.

[9] D. Bates, M. Mächler, B. Bolker, and S. Walker. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*, 2014.

[10] D. Baur. The death of interactive infographics? Blog, Medium.com, March 2017.

[11] E. Bertini and M. Stefaner. NYT Graphics and D3 with Mike Bostock and Shan Carter. Data Stories Podcast, 2013.

[12] B. S. Bloom, M. Englehart, E. J. Furst, W. H. Hill, and D. R. Krathwohl. *Taxonomy of educational objectives*. David McKay Co Inc, 1956.

[13] 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, 2012.

[14] 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, 2016.

[15] 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.

[16] J. Boy, F. Detienne, and J.-D. Fekete. Storytelling in information visualizations: Does it engage users to explore data? In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 1449–1458. ACM, ACM, 2015.

[17] P.-C. Bürkner. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1):1–28, 2016.

[18] S. K. Card, J. Mackinlay, and B. Schneiderman. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.

[19] B. Casselman, M. Conlen, and R. Fischer-Baum. *Gun Deaths In America*, 2016 (accessed October 24, 2017).

[20] M. Feng, E. Peck, and L. Harrison. Patterns and pace: Quantifying diverse exploration behavior with visualizations on the web. *IEEE transactions on visualization and computer graphics*, 25(1):501–511, 2019.

[21] N. Gershon and W. Page. What storytelling can do for information visualization. *Communications of the ACM*, 44(8):31–37, 2001.

[22] 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*, pp. 1191–1200. ACM, ACM, 2015.

[23] Harris, Rich and Popovich, Nadja and Powell, Kenton. *Watch how the measles outbreak spreads when kids get vaccinated - and when they don't*, 2015 (accessed November 2, 2017).

[24] J. Hullman and N. Diakopoulos. Visualization Ehetoric: Framing Effects in Narrative Visualization. *IEEE transactions on visualization and computer graphics*, 17(12):2231–2240, 2011.

[25] J. Hullman, S. Drucker, N. H. Riche, B. Lee, D. Fisher, and E. Adar. A deeper understanding of sequence in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2406–2415, 2013.

[26] M. Kay, G. L. Nelson, and E. B. Hekler. Researcher-centered design of statistics: Why Bayesian statistics better fit the culture and incentives of HCI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 4521–4532. ACM, 2016.

[27] Y.-S. Kim, K. Reinecke, and J. Hullman. Explaining the gap: Visualizing one's predictions improves recall and comprehension of data. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1375–1386. ACM, 2017.

[28] Y.-S. Kim, K. Reinecke, and J. Hullman. Data through others' eyes: The impact of visualizing others' expectations on visualization interpretation. *IEEE transactions on visualization and computer graphics*, 24(1):760–769, 2018.

[29] Y.-S. Kim, L. A. Walls, P. Krafft, and J. Hullman. A bayesian cognition approach to improve data visualization. *arXiv preprint arXiv:1901.02949*, 2019.

[30] S. A. Kluberg, D. P. McGinnis, Y. Hswen, M. S. Majumder, M. Santillana, and J. S. Brownstein. County-level assessment of United States kindergarten vaccination rates for measles mumps rubella (MMR) for the 2014–2015 school year. *Vaccine*, 35(47):6444–6450, 2017.

[31] R. Kosara and J. Mackinlay. Storytelling: The next step for visualization. *Computer*, 46(5):44–50, 2013.

[32] B. Lee, N. H. Riche, P. Isenberg, and S. Carpendale. More than telling a story: Transforming data into visually shared stories. *IEEE computer graphics and applications*, 35(5):84–90, 2015.

[33] M. Lima. Manuel Lima, Information Visualization Manifesto. www.visualcomplexity.com, March 2009.

[34] Z. Liu and J. Stasko. Mental models, visual reasoning and interaction in information visualization: A top-down perspective. *IEEE transactions on visualization and computer graphics*, 16(6):999–1008, 2010.

[35] N. Mahyar, S.-H. Kim, and B. C. Kwon. Towards a taxonomy for evaluating user engagement in information visualization. In *Workshop on Personal Visualization: Exploring Everyday Life*, vol. 3, p. 2, 2015.

[36] L. Masud, F. Valsecchi, P. Ciuccarelli, D. Ricci, and G. Caviglia. From Data to Knowledge - Visualizations as transformation processes within the Data-Information-Knowledge continuum. In *2010 14th International Conference Information Visualisation*, pp. 445–449. IEEE, 2010.

[37] L. Meisler. *Americans Die Younger Despite Spending the Most on Health Care*, 2017 (accessed January 12, 2018).

[38] T. Munzner. *Visualization analysis and design*, chap. 5, pp. 85–107. CRC press, 2014.

[39] Z. Pousman, J. Stasko, and M. Mateas. Casual information visualization: Depictions of data in everyday life. *IEEE transactions on visualization and computer graphics*, 13(6):1145–1152, 2007.

[40] E. Roston, B. Migliozzi, J. Dash, Y. Zhang, B. Migliozzi, A. Pearce, and M. Rojanasakul. *Bloomberg Carbon Clock: Measuring Carbon Dioxide that Causes Global Warming*, 2015 (accessed October 24, 2017).

[41] B. Saket, A. Endert, and J. Stasko. Beyond usability and performance: A review of user experience-focused evaluations in visualization. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization*, pp. 133–142. ACM, 2016.

[42] P. Saraiya, C. North, and K. Duca. An insight-based methodology for evaluating bioinformatics visualizations. *IEEE transactions on visualization and computer graphics*, 11(4):443–456, 2005.

[43] P. Saraiya, C. North, V. Lam, and K. A. Duca. An insight-based longitudinal study of visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1511–1522, 2006.

[44] E. Segel and J. Heer. Narrative Visualization: Telling Stories with Data. *IEEE transactions on visualization and computer graphics*, 16(6):1139–1148, 2010.

[45] M. Stabe. Why the FT creates so few clickable graphics. Financial Times website, 2016.

[46] A. Thudt, J. Walny, T. Gschwandtner, J. Dykes, and J. Stasko. Exploration and Explanation in Data-Driven Storytelling. In N. H. Riche, C. Hurter, N. Diakopoulos, and S. Carpendale, eds., *Data-Driven Sotrytelling*, pp. 59–83. CRC Press, 2018.

[47] A. Tse. Why We Are Doing Fewer Interactives. Malofiej24 Presentation, 2016.

[48] E. Wall, L. M. Blaha, L. Franklin, and A. Endert. Warning, bias may occur: A proposed approach to detecting cognitive bias in interactive visual analytics. In *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 104–115. IEEE, 2017.

[49] W. Wojtkowski and W. G. Wojtkowski. Storytelling: its role in information visualization. In *European Systems Science Congress*, vol. 5. Citeseer, Citeseer, 2002.

[50] J. S. Yi, Y. ah Kang, J. T. Stasko, J. A. Jacko, et al. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization & Computer Graphics*, 13(6):1224–1231, 2007.

[51] E. Zgraggen, Z. Zhao, R. Zeleznik, and T. Kraska. Investigating the effect of the multiple comparisons problem in visual analysis. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 479. ACM, 2018.

[52] C. Ziemkiewicz and R. Kosara. Embedding Information Visualization within Visual Representation. In Z. W. Ras and W. Ribarsky, eds., *Ad-*

vances in Information and Intelligent Systems, pp. 307–326. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. doi: 10.1007/978-3-642-04141-9_15