Brooke Leave Home: Designing a Personalized Film to Support Public Engagement with Open Data

Shauna Concannon†, Natasha Rajan, Parthiv Shah, Davy Smith, Marian Ursu, Jonathan Hook
Digital Creativity Labs, University of York, UK
†CRASSH, University of Cambridge, UK
sjc299@cam.ac.uk, jonathan.hook@york.ac.uk

ABSTRACT
Brooke Leave Home is a personalized film designed to engage a non-expert audience with open data about the support young adults receive when leaving the care system in England. The film draws upon a range of video-based data storytelling techniques to present each viewer with a personalized perspective on the topic based on data from their own local area. We present the film's design and describe how its storytelling techniques were developed to support viewers in understanding, and fostering empathic connections with, the data sources featured and the implications they have for care leavers. We also present a study with 47 viewers, which explores how these techniques were experienced and how effective they were in aiding engagement with the data included and its meaning.

Author Keywords
Data; film; narrative; personalization; storytelling; video.

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

INTRODUCTION
The open data movement has seen increasing amounts of data published by local and national governments, relating to a wide range of societal concerns from funding and tax levels to service performance [7, 34]. Aims of publishing such data include helping to make services more efficient, and improving transparency and [47] involvement in decision making [43]. However, it is widely acknowledged that far fewer members of the public are engaging with available open data sets than intended [46]. Janssen et al. attribute this to challenges including: resources, expertise, capabilities and, perhaps most crucially, motivation to engage using current forms of presentation [36]. As such, much of the public remain unable or uninspired to explore, understand and make use of open data.

One approach to motivating and enabling broader public engagement with data might be to employ new forms of responsive video technology (e.g. [4, 54]) to present data that is personal, or personally relevant, to viewers in video stories. As an example, consider a television news feature that illustrates the impact of national spending changes on a viewer's street by including open government data from their local area. It has been argued that such content could draw on video's well-honed storytelling techniques to provide "easily understood and remembered paths through complex, expansive and conflicting sources of information" and build on the significant popularity of video-based entertainment to present data in forms large, demographically-diverse audiences will be motivated to use [32]. There appears to be much potential in this vision, with nascent examples demonstrating how the creative use of responsive video techniques can lead to accessible, thought-provoking and highly enjoyable experiences of data [1, 26, 28, 62] and related efforts to make data accessible and engaging using other popular storytelling forms, in particular data journalism, proving highly successful [8, 31, 52, 67].

Despite this potential, the design of personalized data videos remains an under-explored challenge. Existing examples only explore a small set of possible techniques for integrating personalized data into video stories, and an established grammar for how such techniques should be used in response to different data sources, messages, production goals and genres is yet to be developed [32]. Moreover, few studies have explored how audiences experience different techniques for presenting data in personalized videos and, perhaps most importantly, whether such videos are actually effective in helping the public to engage with data [3, 32]. As a result, those wishing to make such content lack the evidence base needed to justify commissioning and inform design. This is in contrast to the increasingly mature understanding of narrative techniques and user-experience that exists for other types of data story [31].

In this paper, we present Brooke Leave Home, a personalized film designed to help non-expert viewers understand data about the support young adults receive when leaving the care system in England. We first describe the design of the film and the storytelling techniques it employs to present data and their implications to viewers. We then present a study of how these techniques were experienced by audience members and, crucially, how effective they were in aiding viewers to understand the data and its meaning in the context of the film's topic. Our findings demonstrate how an evocative character-driven video story, combined with the opportunity to explore data by re-watching from different locations, can be highly effective in fostering understanding and empathic connections with data sources and, more generally, how making data understandable, and making people care about data, can be supported through narration, locality, and emotional closeness.

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
http://dx.doi.org/10.1145/3313831.3376462

RELATED WORK

As our lives become ever more permeated by data, developing new forms of data engagement has become a growing HCI concern [18, 23, 49, 50, 53]. The emergence of human-data-interaction as a topic [16, 41, 44] in particular emphasizes a growing focus on goals such as supporting "human manipulation, analysis, and sense-making" of data, "delivering personalized, context-aware, and understandable data from big datasets" and "providing access and understandings of data that is about individuals and how it affects them" [20].

One promising approach for supporting more accessible and appealing human-data-interactions is to present data in media, entertainment and cultural forms people are already familiar with and enjoy. Examples of this strategy include presenting data in visual [56, 66, 68], audio [48, 55] and physical artworks [71] and souvenirs [45, 65]; allowing users to explore and investigate data while playing games [67]; and even those building on contemporary internet culture in the form of memes [19]. Others build on an arts-inspired approach by using aesthetically compelling and evocative representations of data and metaphors that resonate strongly with the user to stimulate engagement and support understanding [10, 12, 21].

A prominent variation on this approach is to use and extend narrative techniques from established storytelling forms to present data – a practice known as data storytelling. Gershon and Page have argued techniques from the "ancient art of storytelling" can help users more effectively understand complex data sources and their meaning [27] and researchers have demonstrated how a variety of story forms can not only support the effective understanding of data, but also make engagement with it fun, fascinating, empathic and even scary or challenging [8, 13, 31, 35, 51, 52]. The potential of data storytelling is perhaps best illustrated by the success of data journalism, where a range of high profile examples (including a Pulitzer Prize winner [40]) show how the coming together of narrative techniques from journalism with data presentation can entice and enable large public audiences to engage [52].

Video is underrepresented in data storytelling research, with comparatively fewer projects exploring it than other forms [1, 3, 2, 32]. This is surprising because video has many properties that could make it an extremely powerful means to bring data to people in ways that they can, and will want to, consume. There are few types of story as approachable, illuminating and entertaining as a professionally created video. Since the first moving pictures in the late 19th century, filmmakers have developed a repertoire of structures, devices and techniques that guide audiences through complex, expansive and often contradictory evidence bases [69]. Furthermore, videos are not only a proven mechanism for explaining complex concepts, but also offer a powerful way to make them evocative, entertaining and enjoyable to consume (e.g., consider how documentaries can make information inspiring, fascinating and critically challenging [6]). A further opportune feature of video is its substantial popularity, as evidenced by viewing figures reported by online video (e.g. YouTube has over one billion users and hours of video watched daily [70]) and streaming platforms (e.g. Netflix alone reports 139 million users [15]). This popularity, combined with the potential of video's diverse genres to appeal to different groups, offers a significant opportunity to create experiences of data that align with the everyday practices and motivations of many people.

As a medium for presenting data to audiences video has a limitation, which may explain why it remains under-explored in this context. Traditional videos are represented as a series of frames that are not easily changed from a semantic perspective and are distributed via a broadcast mechanism. This means data stories based on traditional videos are limited to featuring general data and interpretation, chosen with an entire audience in mind. As a result, the level of individuality that can be presented in terms of data choice and interpretation is limited compared to many of the examples mentioned in the previous section. We observe this might be particularly problematic in the context of experiences for non-expert audiences, as previous work has shown creating personal connections between data and things users care about can be effective in supporting their engagement [42, 33, 30, 24, 37, 63, 22]. Moreover, balancing explanation through narrative and the opportunity to explore based on one's personal perspective has been highlighted as a key design principle of data storytelling [64] – a balance videos identical to all viewers may not strike.

Recent developments in responsive video technology, in particular object-based media [4, 54], allow videos to be dynamically personalized based on data sources. As a result, it is now technically possible to create videos featuring data and interpretation personal to each viewer. While previous work has speculatively argued that this might be used to help people understand data [32], existing empirical research has only explored how a small subset of video's features (e.g. visual narrative, animation) can be employed to this end [3]. We extend prior work by evaluating how a range of further narrative features – including those shown to be valuable in other forms of data story, such as personalization and character [9] – can be used in data videos. In doing so, we contribute a video-based perspective to the constellation of approaches being developed in the CHI community that take data to people in ways that are personally relevant, interesting and accessible [23, 39, 57].

BROOKE LEAVE HOME

Brooke Leave Home is a personalized, data-driven film designed to inform non-expert viewers about the challenges care leavers face in England, and how their ability to tackle them depends on support that varies across local authority areas. Care leavers can be defined as adults who have previously spent time in care (e.g. foster or residential homes) provided by the state or private sector [58]. Those brought up in care often do not have family or other connections to rely on when reaching adulthood [59], so many depend on support from the state (e.g. grants) at this life stage. In England, the film's setting, support provided to care leavers can vary by local authority area. This means the ability of care leavers to face challenges associated with the transition to adulthood, such as setting up a safe and secure home and finding employment, can depend on where they live [59, 60]. This challenge, often referred to as a "postcode lottery" [17], has the potential to be richly illustrated using data about support provision made

ID	Description	Source
1	Care leavers in England by local authority	Children's Society
2	Setting up home grant amount received by care leavers by local authority	Children and Young People Now
3	Estimated care leaver setting up home cost based on itemized basket of goods	Care Leavers' Foundation
4	Availability of council tax exemption for care leavers under 25 by local authority	Children's Society
5	Council tax rates (band D) by local authority	UK Government

Table 1. Data sets featured in Brooke Leave Home.

openly available by governments, or that could reasonably be expected to be made available in the future under common open data policies. Our film aims to raise awareness of this data to audiences and, thus, increase understanding of the challenges care leavers face and promote change.

The film tells the story of Brooke, a fictional care leaver as she begins life on her own in the months following her 18th birthday. The plot follows three acts, broadly serving the purposes of *exposition*, *rising action* and *resolution* of a traditional three-act structure. In the first act, Brooke attempts to set up and furnish her first home using a grant from her local council and a small amount of savings. In the second, she attempts to secure meaningful employment, facing a number of challenges (e.g. debt, a broken washing machine) while searching for jobs. The third act closes the story by showing Brooke's situation after the job search and alluding to future prospects and challenges.

A key feature of the film is that, while all viewers see a story progressing through the same acts, the content and narrative structure shown within them changes depending on data about care leaver support in the viewer's local authority area. The data upon which these differences are based is summarized in Table 1. We intended that this approach would enable us to develop a story that would: i) assist viewers in understanding how complex interactions between multiple data sets affect care leavers lives by using film storytelling techniques to structure their presentation and explanation, in particular a character's story as a unifying narrative path; ii) provide an interesting and evocative human perspective on the data, by illustrating the potential impacts it can have on a character's ability to cope with a series of situations; and iii) enhance the viewer's interest and empathy with the issues in the film, by illustrating how they relate to data that is from their locality rather than a generic location.

In order to ensure that the the character of Brooke, her story and its relationship with the data sources included were constructed in an authentic way, we collaborated with a local not-for-profit social enterprise with expertise in working with care leavers. Staff provided key documents and information sources and, over the course of 8 meetings, offered guidance and feedback on the story. The organization also assisted in arranging and hosting two sessions with a group of care experienced young people, which enabled us to collect first-hand stories that fed into script development. The film was implemented as a web-based experience using the Cutting Room object-based media toolkit [54]. Cutting Room did not support narrative choices based on external data sets and the data-driven graphical elements described in the following sections (i.e. Dynamic Intertitles, Bank Balance Indicator, Re-watch Map). Therefore, these features were implemented using custom HTML, CSS, JavaScript and, in the case of the Re-watch Map, D3.js. The viewer's location was determined from their IP address using the ipapi.co API and the post-codes.co API was used to look up a local authority code. This code provided an index into the film's other data sets, which were stored as .csv files on the film's web server.

In the following sections, we describe the film's design in detail and discuss how the storytelling techniques it employs for presenting data were chosen in response to our aims. We describe the film's main design features as they first appear in the film, but note that each can appear multiple times through the course of the narrative. Three versions chosen to illustrate the key ways the narrative changes in response to data from different locations can be found in the video figure.

Dynamic Intertitles

The film begins with a sequence introducing the viewer to the topic and context of the film, and Brooke as a character. A sequence of shots show Brooke walking through an urban environment, which are punctuated with intertitles (i.e. text shown in between photographed action) that each make a statement from the perspective of a third person narrator. These include: a definition of what a care leaver is and an explanation that they must "stand on their own two feet" immediately after turning 18; a statement of the number of care leavers in England and in the viewer's local authority area; a statement that "The support care leavers receive is different in each local authority"; and an explanation that Brooke is a care leaver setting up home in the viewer's local area after turning 18.

While some of the intertitles are the same for all viewers, two are dynamic and change to include data specific to the viewer's location. The statement of the number of care leavers in the viewer's local area is made personal by including a figure dynamically sourced from a Children's Society data set (Table 1, data set 1). This statement, and one that introduces Brooke and her situation, are personalized by including the name of the viewer's area. For example, if the viewer is in South Norfolk they would see: "This is a story about Brooke, a care leaver who's just turned 18 and is starting out on her own in South Norfolk". Our intent to include data directly in these intertitles wasn't solely to present its content to the viewer (i.e. to make them aware of the potentially large numbers of care leavers living in their area). Rather, we also aimed that presenting raw data values so explicitly at the start of the film would begin to raise awareness among viewers that the film was personalized to their area and data about it, which we felt would be crucial for them to understand some of the less obvious ways the film is personalized to data later on.

Adaptive Voice-over and Shot Choice

In England, the government has a legal responsibility to provide a Setting Up Home Grant that supports care leavers in lieu of the support that a parent might provide (e.g. gifts or loans of household items or money to buy them). The amount provided can vary between local authority areas, and this has been argued to make the transition to independent living more difficult depending where a care leaver lives [17].

The next stage of the film begins with a dynamic intertitle stating the amount Brooke would receive in the viewer's area, which is based on data from Children and Young People Now (Table 1, data set 2). Brooke is shown moving into a new flat and purchasing household items using this grant amount, plus a small amount of savings. This is illustrated using two key shots. First, Brooke is shown flicking through a catalogue in the first person, and speaking about items that she can and can't afford. In areas where the grant is larger, she describes being able to afford items including a new washing machine, a laptop computer and even some luxuries like speakers or a TV license. In areas where it is smaller, she speaks about only buying essentials (e.g. cutlery) and settling for second hand white goods against her support worker's advice. The second shot is a time-lapse of Brooke placing what she's bought into her home. In areas where the grant is larger, the house is shown more completely furnished and the items are newer. The choice about how these two scenes are configured (i.e. what she is shown to be able to afford or not) is made by comparing the amount of money Brooke has with data from The Care Leavers' Foundation (Table 1, data set 3), about common items care leavers must buy when setting up a new home and how much they cost at two retailers.

In presenting scenes showing what Brooke can and can't afford with the grant in the viewer's area, we aimed to evocatively illustrate the implications the grant's value can have on a care leaver's ability to meet the challenge of setting up home. The scenes also serve a purpose in foreshadowing a range of challenges that can arise in act two as a consequence of Brooke's purchasing ability. The second scene closes with a shot of Brooke playing the guitar in her new home and speaking about the importance of music in her life – establishing a key aspect of her character that shapes her decisions later in the film.

Bank Balance Indicator

When creating the previous scenes, we were keen to help viewers understand that while a Setting Up Home Grant can look quite large when expressed as a raw number (e.g. ranging from £750 to £2,500 in the data set used), this amount is quickly spent when all the items required to set up a home are considered. When viewing initial edits of these scenes as described above, we found they did a good job of conveying the major purchasing decisions that Brooke would have to make differently based on the grant available to her. However, we were concerned that they did not provide an adequate enough sense that a large percentage of her grant would be spent on a long list of low-value, everyday items (e.g. light bulbs, a mop). This, interestingly, was in contrast to viewing the raw data about the items for purchase, which provided a vivid illustration of this point due to its substantial length.

In response to these concerns, we added a dynamic graphical overlay to these, and subsequent, scenes. This graphic, the Bank Balance Indicator (Figure 1), shows how much money Brooke currently has in the bottom-right corner of the screen. When Brooke refers to making purchases in her voice-over, the balance is reduced and the items bought and their price are shown as a list of small labels above this value. These labels fade out over the course of a few seconds following a

Figure 1. Bank Balance Indicator showing purchases being deducted.

purchase. We use the Bank Balance Indicator in two ways in the previously described scenes. Firstly, we use it to show the cost of major purchases mentioned by Brooke (e.g. her washing machine), deducting these as they are referred to in the voice-over. We intended that showing data in this way would establish a degree of authenticity in the editorial decisions made by showing the data they are based upon. Secondly, we aimed to convey that Brooke's budget is quickly depleted by a large number of small purchases, by rapidly deducting all of the other items included in the data set over the course of a short section of voice-over where Brooke talks about having to buy "absolutely everything". Our intent in displaying the data in this way wasn't to allow the viewer to read and interpret all items. Rather, it was the opposite: to convey a feeling that Brooke is overwhelmed by unexpected, small expenses, by showing them faster than they can be read.

Complicating Factor Scenes

The second act follows Brooke as she seeks her first job. We employ an adaptive narrative structure in which Brooke's ability to meet the challenge of searching for a job, preparing a CV and attending an interview are shown to be affected by a number of complicating factors. The factors shown are chosen based on aspects of Brooke's situation established in the first act, as well as further data about the support available in the viewer's local area introduced as the story unfolds.

The first factor that may complicate Brooke's job search is being in tax arrears. In some areas of England, care leavers under 25 are exempt from paying tax to their local authority, while in others they are not. The Children's Society have argued that this additional, often unexpected, expense can have a significant negative impact on care leavers, many of whom are already under financial pressure and have little experience of managing their money [60]. In the film, Brooke is shown searching the internet having received her first council tax bill. She's shown to be in distress, as she hadn't factored this expense into her planning. In areas where Brooke is exempt as a care leaver, she is shown finding this out and expressing relief. In areas where she isn't, she is shown to be visibly worried and states her job search is more urgent than it was before. Additionally, the amount of council tax she owes is deducted from the Bank Balance Indicator. These narrative choices are based on data from the Children's Society about which English councils offer a tax exemption for care

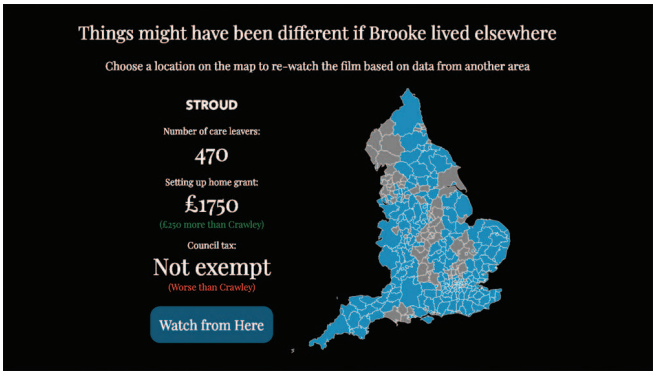

Figure 2. Map for browsing data and choosing re-watch location.

leavers (Table 1, data set 4) and data from the UK Government detailing council tax rates by area (Table 1, data set 5). The scene closes with an intertitle that explains that council tax is not paid by all care leavers in England, and shows the amount that Brooke either has to pay or is exempt from.

As the job search continues, a number of further complicating factor scenes are shown or not depending on Brooke's situation. If Brooke was able to afford a laptop computer in act 1, she is shown using this to conduct a productive job search and prepare a CV. However, if she wasn't, she's shown struggling to do this before the allotted time on a shared computer in her local library runs out. If Brooke is successful in preparing her CV, but finds herself under pressure to get a job quickly due to being in debt, she's shown spending valuable time that could be spent searching for her ideal career working as a cashier in a cafe. This work provides a small amount in wages that are reflected on the Bank Balance Indicator. If Brooke was only able to afford a second hand washing machine in act 1, this breaks down and she's shown struggling to wash her clothes in the bath and feeling worried about attending interviews in dirty clothes. If Brooke is still in debt in the later stages of her job search, she's shown selling her most prized possession, her guitar. Finally, if Brooke remains in a large amount of debt after her guitar has been sold, she is shown struggling to pay the bills and afford food. Our aim in adopting this structure was to further illustrate the human consequences that can result from the data, by showing how they might directly (e.g. council tax debt) or more obliquely (e.g. broken washing machine) frustrate a care leaver's ability to succeed.

Ending Based on Character State

The final act shows the outcome of the job search and looks forward to the future in three possible endings. In the first, Brooke has had space to think about what she'd like to do with her life, and time to prepare. She follows her passion for music, securing a role with a charity and recording her songs as part of a project. In the second, Brooke has been distracted by a sub-set of the complicating factors in act 2, and as a consequence has abandoned her plans for a career in music. Instead, she's shown working as a cashier in a cafe, with the situation unlikely to change. In the final outcome, Brooke is highly distressed by the coming together of multiple complicating factors. She's unable to secure employment and is shown surrounded by bills as bailiffs knock at the door.

The ending is chosen based on a variable representing Brooke's level of preparedness and mental state following the challenges faced. This variable is incremented or decremented during act 2 depending on the scenes shown. E.g., it is incremented if Brooke uses a laptop to properly prepare her CV, but decremented if she is distracted by taking on part-time work or caused distress by mounting financial pressure. Our aim in using a single variable, updated in response to multiple data sources, to represent Brooke's state at the closing of the film was to allow an ending to be chosen based on the ways these data sources can interact to impact her life.

Re-watch Structure and Map

While a central aim of the film was to provide viewers with an accessible and interesting representation of data that is personally relevant to them, we were also mindful that its core message (i.e. support for care leavers across England is uneven) hinged on the comparison of this personalized data and its implications with other areas. To support this, we implemented a storytelling technique where viewers are able to *re-watch* the film from multiple locations. The film ends with two intertitles, one that states that Brooke's story was based on data from the viewer's local area and a second that asks "Would things be different if Brooke lived somewhere else?". The viewer is then presented with an interactive map of England (Figure 2). To the left of this map is a panel that shows the raw data values that the film is based upon. The viewer can click on different locations on the map, to be shown the data from that local authority area and how it compares to the data from where they've just seen the film. The viewer can click "Watch from Here" and the film is shown again with the content and narrative based upon data from the selected area.

We aimed that this technique would allow viewers to understand the broader context of the data from their area in two ways. Firstly, we intended the map would provide a lightweight way for viewers to quickly view data from multiple local authority areas and, thus, gain a broader understanding of the divergent levels of support across England. Secondly, we aimed that once viewers had browsed the map and decided on an area to re-watch the film from, they would be able to gain a clear and evocative picture of what these differences can mean for care leavers, by seeing how they change Brooke's ability to cope with the same situations previously seen.

STUDY DESIGN

To evaluate Brooke Leave Home we conducted an online study in which viewers completed a pre-study survey, watched the film, and then completed a post-study questionnaire about their viewing experience. The study was designed to evaluate responses to the storytelling techniques employed, with a focus on evaluating how effective they were in providing an engaging and easily understandable account of the data featured. In particular, we wanted to investigate the effect of two key design choices and assess: if the re-watch functionality facilitated a better comprehension of the topic at a national level and provided a fuller understanding of the topic; and whether viewers engaged more with the film if it was set in, and incorporated data related to, their local area. To explore the effect of these particular design features, each viewer was

assigned at random to one of four conditions: i) local version, re-watch enabled (LR); ii) local version, no re-watch (LNR); iii) non-local, re-watch enabled (NLR); iv) non-local, no re-watch (NLNR). Viewers in local conditions saw a version of the film using data from their local area, while those in the non-local conditions saw a version based on data from another randomly selected area. Viewers in the re-watch conditions were presented with the Re-watch Map after viewing the film, while those in non-re-watch conditions were not.

Apparatus

The experiment detected the viewer's location as a local authority code, ensuring those in local conditions saw a film based on data relating to their local authority, and viewers in the non-local conditions did not. When it was not possible to automatically detect a user's location (e.g. due to antivirus software causing the location request to fail) participants were shown a version of the film from a randomly selected location and automatically assigned to a non-local experiment condition (i.e. NLNR or NLR) to reflect this. We subsequently checked all non-local viewing locations against hometown data provided in the pre-survey to ensure none were randomly allocated their actual location in error. Each viewer received a unique link to a version of the film online to be watched in a web browser. We collected a log of details relating to the particular version of the film seen by each viewer, i.e. the location the film was based in, the location specific data it used and the scenes selected for the narrative, including the ending. The data was logged in an SQL database. The pre-study survey and post-viewing questionnaire were accessed via online forms hosted on the university's secure server.

Participants & Materials

Viewers were recruited via a number of channels, resulting in a sample comprising university students, staff and general public. 54 viewers started the study; 7 dropped out after the pre-study survey and were excluded from analysis. A total of 47 completed the study and received a £5 Amazon voucher for their time. The viewers covered a range of ages: 38.3% (n=18) aged 18-24, 36.2% (n=17) aged 25-34, 19.1% (n=9) aged 35-44, 4.3% (n=2) aged 45-54 and 2.1% (n=1) aged 55-64 years. All viewers were resident in England. The automatic re-classification of conditions when location requests failed resulted in more viewers being placed into non-local conditions, leading to some inequality of sample size (LR n=11, LNR n=10 , NLR n=13 and NLNR n=13).

The personalized form of Brooke Leave Home meant that the film presented to viewers varied in a number of ways. We tried to recruit viewers from different locations to cover the range of data possibilities that the film could represent. Due to the nature of the study, it was not possible to entirely control for the distribution of locations and related data points. A key difference in the viewers' experience of the film is shaped by the narrative trajectory taken, and the associated endings. 12 unique paths through the film's scene structure were seen by viewers in the study. Of the three possible endings, 17 viewers (on first watch) saw the most positive, 19 the moderate outcome and 11 the least positive.

The pre-study survey gathered contextual information about each viewer including: age bracket; whether they had been a resident in England for more or less than 5 years; the first part of the postcode for their current location (to confirm the part of the country they are in); and places they have a strong association with in the UK, e.g. hometown (to see if re-watch selection was impacted by a personal connection to a place). The post-film questionnaire gathered qualitative and quantitative data about the viewers' response to the film, in terms of i) general understanding of the topic, as well as more specifically, how the support for care leavers varies nationally; ii) engagement with the character-led story and empathic response; and iii) reflections on use and display of data in the film. Viewers in the re-watch conditions were asked additional questions to ascertain how they interacted with this design feature and details of their experience of re-watching the film. Quantitative comparisons of understanding and empathy in our analysis are based on self-reported measures of topic awareness and how much each participant empathised with the character reported before and after watching on a scale of 1-5. Qualitative comparisons of empathy are based on coding excepts for self-reports of empathy to Brooke or wording suggesting an appreciation of her feelings [11].

RESULTS

Descriptive Statistics

Participants provided a series of ratings, recorded on a 5-point Likert scale. Unless otherwise stated the point scale was 1=*Not at all* to 5=*Very much*. On average, viewers were not overly familiar with the topic of the film before watching ($\overline{X} = 1.89$, SD: 1.05). Viewers across all conditions generally found the film easy to understand ($\overline{X} = 4.53$, SD: 0.86, 1=*Difficult*, 5=*Very easy*) and interesting to watch ($\overline{X} = 4.09$, SD: 0.86). The film was rated positively for relevance to the viewer's society and local community ($\overline{X} = 4.40$, SD:0.74). The average rating for the perceived objectivity of the film was less conclusive, but still generally positive ($\overline{X} = 3.28$, SD: 0.99). Viewers across conditions rated that they felt a strong empathic response to the character of Brooke ($\overline{X} = 4.28$, SD: 0.85), and felt more concerned than hopeful about the experiences of care leavers ($\overline{X} = 1.94$, SD: 1.09, 1=*Concerned*, 5=*Hopeful*).

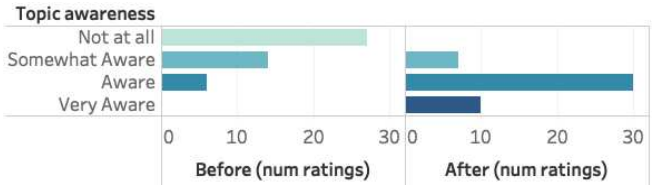

Figure 3. Aggregate awareness ratings before and after watching

Awareness and Understanding

55.32% of viewers (n=26) explicitly reported the film influencing their opinions or feelings on the topic. To assess the shift in opinions and attitudes, viewers were asked to provide a rating of their awareness of the topic and its importance before and after viewing. A Wilcoxon Signed-Ranks Test indicated that the median post-watch awareness levels were statistically significantly higher than the median pre-watch ranks ($\tilde{Z}$ 5.78, $p < 0.001$), as were importance ratings ($\tilde{Z}$ 5.20, $p < 0.001$).

Figure 3 details the total number of viewers who rated their awareness of the topic, both before and after watching.

Bank Balance Indicator

42.6% of viewers (n=20) reported that showing Brooke's purchases on the Bank Balance Indicator strengthened the connection between the fictional account and the underpinning data *Very much*, 36.2% (n=17) *somewhat* and 21.3% (n=10) *Not at all*. On the extent to which it influenced how fact-based the story was, 76.6% respondents answered *Very much*, 21.3% answered *Somewhat*, and 2.1% (n=1) *Not at all*, with a similar distribution reflected in response to how it affected the stories authenticity: 74.5 % (n=35) of viewers answered *Very much*, 23.4 % (n=11) reported "Somewhat" and 2.1 % (n=1), "Not at all". Figure 4 shows the distribution of these responses.

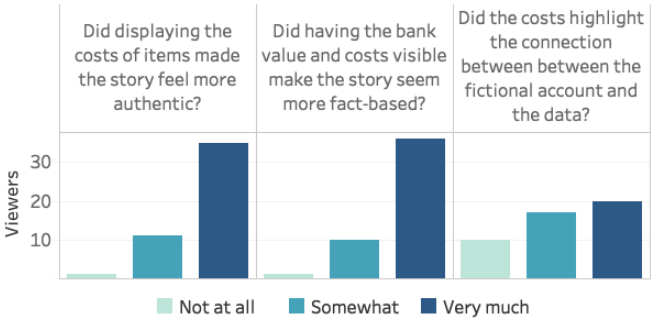

Figure 4. Responses to questions about the Bank Balance Indicator

Use of Re-watch Functionality

There were 25 viewers in conditions with re-watch functionality (LR, NLR). Only one viewer didn't engage with the re-watch functionality at all, with the remaining 24 re-watching the film between 1 and 11 times. The mean number of re-watches per viewer was 2.76 (SD: 2.33). The mean number of interactions with the map per viewer (i.e. locations clicked) was 22.20 (SD: 20.04), with a range from 0-76.

All viewers were asked to name their current location and home town in the pre-study survey. For viewers in non-local conditions, we compared these survey responses with a manual inspection of the locations explored on the map, and locations from which the film was re-watched. All 13 viewers who had watched a non-local version of the film explored the data for their current location or home town on the map, except for one participant who didn't engage with the map at all. 8 of these viewers also watched the version of the film based on their local data. Those who didn't tended to select locations with data that was distinct and which led to different endings compared to the film they had just watched (e.g. a lower or higher setting up home grant or different tax exemption status).

Evaluating the Design Features

To understand the effect of setting the film in the viewer's local area and the re-watch structure and map had on understanding, engagement and empathy with the topic we tested the effect of condition on viewer responses using non-parametric tests for independent samples, with post-hoc pairwise comparisons included when a general effect of condition was observed. An independent samples Kruskall Wallis Test finds an omnibus effect of condition on the understanding of situation at a local level ($H = 9.048$, $p < 0.05$). However, post-hoc pairwise comparisons, with Bonferroni corrections show no asymptotic significance between conditions. Ratings of empathy with Brooke did not differ significantly across conditions ($p < 0.06$), nor did ratings for how relevant the film was to their local community ($p < 0.90$). This is counter to our hypothesis that viewers in local conditions would feel more empathy for Brooke and the issue more relevant to their local community.

However, a Kruskall Wallis test finds a significant effect of condition on the understanding of the situation at a national level ($H = 25.873$, $p < 0.01$). Post-hoc pairwise comparisons, with Bonferroni corrections show significant effect of conditions in all comparisons between re-watch and non re-watch conditions. The planned pairwise comparisons reveal a significant difference between the LNR and LR conditions ($p < 0.01$), LNR and NLR ($p < 0.01$), LR and NLNR ($p < 0.02$), NLR and NLNR ($p < 0.01$). There was no significant difference between the two conditions which did *not* include the re-watch feature ($p = 1.00$), nor between the local and non-local conditions which *did* include the re-watch feature ($p = 1.00$). This suggests that the re-watch feature facilitates understanding at the national level, which was our design intention. This highlights the importance of being able to compare and contrast data examples to explore what data means in context.

Slightly more viewers in the re-watch conditions reported that they thought care leavers were not given sufficient support (75.00%) than in the non-re-watch conditions (60.87%), although this difference was not significant ($\tilde{\chi}^2$ 1.08, $p < 0.29$). Viewers were more likely to report that care leavers did not receive equal support across the country if they were in the re-watch conditions ($\tilde{\chi}^2$ 11.61, $p < 0.001$). 100% of participants in the re-watch condition, when asked if there is equal support available in different parts of the UK, responded no, compared to only 60.87% of respondents in the non-re-watch conditions.

Qualitative Analysis

An Accessible and Engaging Form of Data Presentation

Viewers found the film presented information in a clear and easily understood manner, describing it as "educational and insightful", "informative", "incredibly accessible", "easy to follow" and "digest" and providing "instant clarity". Viewers were asked to summarize the key message of the film, and demonstrated an ability to articulate it accurately. For example, P41 (LR) reflected that care leavers faced a number of challenges that could be exacerbated depending on the support provided locally: "being a care leaver can be difficult, especially based on where you're living. It isn't always easy to find a job, and when you do it may only be minimum wage - certainly not enough for some people, especially care leavers" and P44 (NLR) highlighted the importance of adequate support for future achievement: "the provision of support for care leavers has a vital impact in helping care leavers achieve a future. Even a small budget can make a big difference".

In reference to new knowledge gained from watching the film, a number of viewers reported gaining a fuller understanding of the situation encountered by care leavers. For example, P47 (NLNR) explained "I didn't realize how many care leavers there were or that they were just thrown into the deep end in terms of having their own home but being given very little

help or money to support themselves". A number of viewers reported how the film aided their understanding of the national variation in support, as P7 (NLNR) summarized, it made "clear the disparity across regions and in the outcomes for young people". P12 (LNR) commented that they now considered "how important funding and support for care leavers is" and how "this potentially varies depending where you live".

16 viewers noted that the film provided a "good insight" on an unfamiliar topic, and a new perspective they "hadn't considered" or "thought about". P58 (NLNR) remarked "I had no idea how [the care system] works", while P25 (NLNR) explained it introduced "the experience of a care leaver when I had never considered their situation before". For participants with some prior awareness of the topic, the film deepened understanding and empathy. E.g., P50 (NLNR) said "I always knew that people leaving the care system were not receiving enough support, but seeing it played out by a real person made me realize it even more, and increased my awareness of just how little help they receive. I now feel even more concerned about these people and angry on their behalf".

Blending Facts and Fiction

Combining factual content with a fictional, although factually informed, story was generally positively received; it "presented the data in a human way that you can relate to and empathize with" (P12 LNR), "effectively illustrat[ing] the importance of money in their lives and futures in a way that just reading data would not" (P27, NLR). Video was seen to have a number of advantages for data presentation, such as being "easier to follow and relate to, more engaging", and P16 (NLR) explained, that "as a visual learner I appreciated this".

The evocative capacity of film was a common theme, with viewers expressing that Brooke Leave Home evoked an emotional response, e.g. "It was emotional watching one young woman's struggles" (P26, NLNR); "It's far more impactful and has an emotional resonance that drives the viewer to empathize with care leavers" (P27 NLR). The emotional connection was particularly anchored in response to the character-driven representation of data, as P5 (LR) reported: "By making it a fictional story, based on the facts but at the same time providing a character to empathize with makes the film evocative for me". Viewers, such as P27 (NLR), demonstrated their empathy for the character: "in a region where Brooke had no money, it really hit me that she had nowhere to turn". The character-driven narrative was a key factor in fostering an understanding of the lived experiences of care leavers and considering how this related to the data. For example, P12 (LNR) suggested, it "presented the data in a human way that you can relate to and empathize with" and P58 (NLNR) that it "[showed] what it means for a human being to live with the given support".

Potential for Oversimplification

The simplicity and character focus of the video story appealed to many viewers, however, some highlighted its tendency for oversimplification of a complex issue, e.g. P27 (NLR) "the form of this data driven video requires over simplification of difficult issues". Furthermore, while the adaptive nature of the film allowed greater flexibility than a traditional film, it is not possible to encapsulate the multiplicitous and varied experiences of all the people the data represents. Prioritizing certain eventualities in the narrative over other possibilities, was identified as a reductive limitation by P9 (LR): "it assumes that certain factors effect peoples thinking in the same way, people are different with varied life experience based on many factors". Some viewers identified other modes as better suited to support data exploration, suggesting that it "may not be as nuanced or detailed as an article or a data visualization" (P12, LNR) or as well suited "for exploring the diversity of provision across the country" (P5, LR).

The potential for oversimplification was seen as a possible detractor for trust. Some viewers felt that while the video story was "more engaging" it can also "appear as less factual" (P44, NLR). As P28 (LNR) observed, "it doesn't necessarily represent a more generalised approach, which could make it harder for people to trust information about broad, national trends". Most viewers reported that the integration of data and fictional account boosted the authenticity of the story, but for others it undermined perceived objectivity. While "giving a name and face to a person going through this makes it more interesting and evokes more sympathy", P25 (NLNR) felt there was "a limit to how much data you can present in this way as I felt the main emphasis was on telling the story from her point of view and therefore was not very objective".

Making the Abstract Concrete

Making abstract figures meaningful was highlighted as a key advantage. By illustrating "very concrete experiences that can result from the abstract financial decisions", P22 (LR) explained, the film was able to "translate abstract figures e.g. home grant amounts, into a more textured understanding of the everyday barriers that those figures might create [...] £2k might sound like a generous setting up home grant. But seeing the trade-offs that the grant amount led Brooke to make, gives a different understanding of how useful that grant really is".

The Bank Balance indicator was particularly effective in drawing the narrative and the data together. P53 (LNR) noted "it was nice to have the information visible. It highlighted how quickly her money ran out" and P9 (LR) found "it helped to visualized that her funds would not go that far". Furthermore, it was said to reinforce how the initial pressures associated with setting up home could impact negatively on life decisions: "seeing the balance go into the negative so quickly really makes it clear that work options are not considered, but work is taken out of necessity, and sadly discourages young care leavers from being able to carve out a career which may prove to be more financially stable in the long run" (P32, NLR). The Bank Balance Indicator and textual display of data on the intertitles, coupled with the video narrative, was also shown to deepen viewers' understanding of what the values meant in real terms, and improve the authenticity of the story. For example, P48 (NLR) reported:"It was factual but still authentic" and P17 (LNR) explained "I liked the written facts and then seeing how she is dealing with those in real time. It makes it more real. Words on a piece of paper seem a little detached from the reality. Give a face or voice to the actual experience makes the point more and connects with you emotionally".

Comparison and Contextualization via Re-watch

Echoing the quantitative findings, re-watching the film from different locations was seen to enable viewers to comprehend the variation of support across the country and understand what these differences could mean in real terms. P41 (LR) noted, "seeing the film actually play out differently rather than just being told Brooke wouldn't be able to get a job (for instance) helps the viewer to realize the issues with the system". The concept of the re-watch appealed to many of the viewers, as it facilitated the direct comparison of different data points in a meaningful way. For example, P62 (LR) explained: "Made it much clearer, especially when watching the film again in a different location [...] seeing how difficult it was in an area where Brooke received even less money [...] makes it easier to understand the experiences of people in your area and facilitates comparisons".

There was an appetite for re-watching the film from other locations to contextualize and draw comparisons, with 6 viewers in the non-re-watch condition expressing a desire to re-watch from another location and 18 indicating they wanted to know more about how support for care leavers varies across the country. Viewers who did not have the opportunity to re-watch the film were concerned that the approach gave them a partial understanding of the topic, as "you don't get to see the comparison between living in different areas" P53 (LNR). P58 (NLNR) wanted to know how the version they watched "compares to other places in the UK? [...] As a viewer, you miss out on other possible perspectives. One story does not describe the variety of ways care leavers start out on their own". It was observed that if you only see one version of the film based on a single location, you "may lose sense of comparison" (P7, NLNR), serving to "emphasize the disparity between the different areas of the UK", (P4, NLNR), giving a "polarized view"(P17, LNR) and creating "a falsely positive (or negative) impression of what it's like for care leavers generally" (P12, LNR). Notably, concerns that the film could misinform or provide a skewed take on reality, were only raised by participants who did not have access to the re-watch functionality, suggesting that re-watching from a different location *did* support a more contextualized understanding of topic and data.

As illustrated by the quantitative data on the re-watch views, individuals did make use of the function. Viewers who were exposed to the re-watch functionality highlighted how it enabled them to draw comparisons in the data: "it was great to see the difference the money actually made to Brooke's life" (P27, NLR). However, although they liked the concept, 8 participants explicitly reported they found the process of re-watching laborious or repetitive: "I would not want to watch it over and over. 2 times was definitely enough" (P1, LR). The time required to re-watch multiple versions of the film was identified as a potential barrier for viewers to achieve the comparative interactions with the data that the re-watch functionality intended to support: "may take longer to watch (than if it were something like a data visualization)" (P44, NLR). This is supported by the re-watch log data, which shows 44.80% of re-watch views were terminated before the film ended. While many viewers found it "impactful to see the same situation turning out a different way" (P27, NLR), 5 viewers commented on the similarities in the different versions. 3 viewers expressed that the duplication of content impacted negatively on their engagement with subsequent watches, as "it looked similar". For example, P5 (LR) remarked: "the different areas do not feel distinctly different since the actor and general visual aesthetic did not change from place to place [...] there isn't much novelty in watching a slightly different version of a video".

Interest in Local Data

While the quantitative analysis showed little effect of condition in relation to localization, a large number of viewers highlighted how it impacted on their response to the film and motivated their location selection for re-watching. For example, P48 (NLR) wanted "to see what its like for care leavers who live near me", and as P32 (NLR) explained, exploring data about your local area can enable you to better connect with or relate to the issue: "it is always interesting to see how people are affected in terms you can relate to". Viewers in the non-local conditions also expressed a desire to see localized content or conveyed that they thought it would have been more powerful: "I would take more notice if it were my local town/city" (P25, NLNR). There was also a desire for more tailored and nuanced localization, by, for example, referencing local initiatives or "a [local] scheme that Brooke would've accessed in reality" (P62, LR).

DISCUSSION

At the start of this paper we hypothesized that personalized video stories can be a powerful way to motivate and enable members of the public to understand open data and its implications. Our quantitative findings show significant increases in ratings of topic awareness and importance after watching; that most viewers found the presentation of data to be interesting, easy to understand and relevant to their local community; and that viewers felt a strong empathic response and concerned about care leavers after watching. Qualitative responses also show the film was accessible and informative and viewers were able to summarize the main messages of the film accurately, suggesting it was effective in promoting understanding. This evidence leads us to conclude that, in the case of Brooke Leave Home, a personalized film was effective in making data understandable and viewers care about it. In order to inform the design of future personalized data videos, we now discuss our findings to explore how the various storytelling techniques employed in the film supported these levels of understanding and empathy observed in our audience.

Character Focus for Understanding and Empathy

The overarching storytelling technique employed in the film was to demonstrate the impact that different data values can have on peoples' lives by illustrating how they affect a character's ability to negotiate a series of challenges. Our qualitative findings suggest that such responsive, character-focused narratives can offer a powerful means to help viewers understand the human implications of data and, in particular, to evoke heightened empathy with the people data affects. Previous research has argued fostering such empathic and emotional connections with data can be important for supporting sensemaking [9, 38], suggesting this aspect of the film's design may

have been particularly integral for the levels of viewer understanding observed. A potential limitation of our approach in this context was highlighted in qualitative responses by 3 viewers, who noted using the same character for multiple viewings might negatively effect engagement. However, a lack of significant differences in level of empathy across the non-rewatch and re-watch conditions suggests this may not be a serious limitation of the technique for future design. When designing the film, we were mindful that the success of our character-focused approach would depend on making the relationship between the underlying data and the narrative structure chosen for each viewer comprehensible. Findings suggest that the presentation of raw data values in, and alongside, visual content (e.g. in the Dynamic Intertitles and Bank Balance Indicator) offered an effective design technique for achieving this goal – with the indicator's synchronized display of data values during scenes noted, in particular, for making the film feel more fact-based and increasing perceived authenticity.

Data Personalization as a Powerful Motivating Factor

We hypothesized that including data and interpretation personal (i.e local) to the viewer would increase their level of interest and empathy and, thereby, enhance understanding. Our quantitative results did not provide any evidence of such an effect. However, qualitative results did show positive viewer responses to the film's localization, which relate to enhancing feelings of personal relevance, familiarity and empathy with the topic. Moreover, findings relating to the re-watch technique show viewers had a strong appetite to view data from their area – with those in the NLR condition predominantly choosing to re-view the film from a local, or other personally significant, location and those in the NLNR condition expressing a desire to do so. These findings suggest that, while personalization was not as powerful of a device for supporting understanding and empathy as we had expected, it can be a particularly effective design technique for motivating viewers to engage with data and its implications.

Re-watching to Contextualize a Personal Perspective

Our intent in designing the re-watch technique was to enable viewers to understand how data from their local area, seen on first viewing, relates to issues facing care leavers at a national scale. Both sets of findings provide evidence the technique achieved this goal, with quantitative comparisons between re-watch and non-rewatch conditions in particular showing significant differences in understanding at a national scale. These results suggest re-watching has the potential to be an effective storytelling technique in cases where a personalized account of data needs to be contextualized in a data video. When designing the technique, we were concerned that viewers would find re-watching tedious. Our findings suggest that viewers were willing to re-watch the film multiple times. However, they also suggest that after 2-3 viewings the experience can become repetitive. In future data videos adopting a re-watch technique, especially longer ones, this issue might be addressed by using responsive video approaches (e.g. [25]) to abridge later viewings by, e.g., removing or shortening scenes or only re-showing aspects that highlight differences. Our findings also show that many viewers used the map to browse data from multiple locations before re-watching from one, with the average number of locations clicked on exceeding the average number of re-watch views. This suggests the map's value was not only in functionally enabling the selection of areas for re-watch – but that it facilitated understanding in its own right, with video viewings stoking curiosity that, in turn, motivated broader engagement with data via the map's alternative, more exploratory, mode of presentation.

Risks of Simplification and Perceived Objectivity

While combining factual and fictional content facilitated an engaged and emotional connection to an otherwise unfamiliar topic, some viewers felt this led to a simplified representation that limited perceived objectivity. These findings highlight potential issues of trust and representational complexity that may arise as data videos seek to aid users in understanding data and its implications by swapping free exploration for greater levels of interpretation from a defined authorial viewpoint – a trade-off that has previously been observed across a range of other data storytelling forms [64]. We believe that – like in traditional documentary film, where the dialectic between objective truth and representation is a well acknowledged concern [14] – this challenge can be negotiated by reflective and ethical practice by authors in respect to their intentions, viewpoints and potential biases. However, we also believe our findings show that negotiating such issues may prove especially complex for content that draws attention to being *data-driven*, in particular due to the well-established status of data as an objective thing in our society. For instance, our findings suggest that design techniques that present data prominently in the course of a video, like our Bank Balance Indicator, can make a story feel more authentic and fact-based. On one hand this finding has positive implications, suggesting that the Bank Balance Indicator and similar techniques can help reinforce viewer trust that an ethical position has been taken. However, on the other, it highlights an amplified risk of misleading audiences about the objectivity of views, either intentionally or not, when using design techniques that present data prominently in videos.

Moving From Awareness to Action

A key motivation for Brooke Leave Home was to increase understanding of data about the challenges care leavers face and, thereby, promote change. While the film was successful in raising awareness, research on citizen participation suggests that this alone will not likely empower citizens to change situations. E.g., Arnstein notes that informing people can be "the most important first step toward legitimate citizen participation" but if not coupled with mechanisms for empowerment can lead to tokenism [5]. For this reason we argue for future work that not only explores the form of data videos, but also how they might be situated in broader structures that enable citizens to change situations. Examples of ways to achieve this might include situating films in charity campaigns (e.g. [61]) or employing iDoc forms that evolve over time via viewer participation to scaffold activist communities [29].

ACKNOWLEDGMENTS

This research was funded by EPSRC grants Perspective Media (EP/R010919/1) and Digital Creativity Labs (EP/M023265/1).

REFERENCES

[1] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Christophe Hurter, and Pourang Irani. 2015. Understanding Data Videos: Looking at Narrative Visualization Through the Cinematography Lens. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1459–1468. DOI: http://dx.doi.org/10.1145/2702123.2702431

[2] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Andrés Monroy-Hernandez, and Pourang Irani. 2017. Authoring Data-Driven Videos with DataClips. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan 2017), 501–510. DOI: http://dx.doi.org/10.1109/TVCG.2016.2598647

[3] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Jason Leboe-McGowan, and Pourang Irani. 2018. Hooked on Data Videos: Assessing the Effect of Animation and Pictographs on Viewer Engagement. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces (AVI '18)*. ACM, New York, NY, USA, Article 21, 9 pages. DOI: http://dx.doi.org/10.1145/3206505.3206552

[4] Mike Armstrong. 2014. Object-based broadcasting - curation, responsiveness and user experience. *IET Conference Proceedings* (January 2014), 12.2–12.2(1). DOI:http://dx.doi.org/10.1049/ib.2014.0038

[5] Sherry R. Arnstein. 1969. A ladder of citizen participation. *Journal of the American Institute of planners* 35, 4 (1969), 216–224. DOI: http://dx.doi.org/10.1080/01944366908977225

[6] Paul Ashton. 2014. *The Calling Card Script: A writer's toolbox for screen, stage and radio*. Bloomsbury Publishing.

[7] Judie Attard, Fabrizio Orlandi, Simon Scerri, and Sören Auer. 2015. A systematic review of open government data initiatives. *Government Information Quarterly* 32, 4 (2015), 399 – 418. DOI: http://dx.doi.org/10.1016/j.giq.2015.07.006

[8] Benjamin Bach, Nathalie Henry Riche, Sheelagh Carpendale, and Hanspeter Pfister. 2017. The Emerging Genre of Data Comics. *IEEE computer graphics and applications* 38, 3 (2017), 6–13. DOI: http://dx.doi.org/10.1109/MCG.2017.33

[9] Benjamin Bach, Moritz Stefaner, Jeremy Boy, Steven Drucker, Lyn Bartram, Jo Wood, Paolo Ciuccarelli, Yuri Engelhardt, Ulrike Koeppen, and Barbara Tversky. 2018. Narrative Design Patterns for Data-Driven Storytelling. In *Data-Driven Storytelling*. AK Peters/CRC Press, 125–152.

[10] Scott Bateman, Regan L. Mandryk, Carl Gutwin, Aaron Genest, David McDine, and Christopher Brooks. 2010. Useful Junk?: The Effects of Visual Embellishment on Comprehension and Memorability of Charts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 2573–2582. DOI: http://dx.doi.org/10.1145/1753326.1753716

[11] C. Daniel Batson. 2009. These things called empathy: eight related but distinct phenomena. In *The social neuroscience of empathy*. MIT press. DOI:http://dx.doi.org/10.7551/mitpress/9780262012973.003.0002

[12] Fadi Botros, Charles Perin, Bon Adriel Aseniero, and Sheelagh Carpendale. 2016. Go and Grow: Mapping Personal Data to a Living Plant. In *Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI '16)*. ACM, New York, NY, USA, 112–119. DOI: http://dx.doi.org/10.1145/2909132.2909267

[13] Jeremy Boy, Francoise Detienne, and Jean-Daniel Fekete. 2015. 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 (CHI '15)*. ACM, New York, NY, USA, 1449–1458. DOI: http://dx.doi.org/10.1145/2702123.2702452

[14] Stella Bruzzi. 2006. *New documentary*. Routledge.

[15] Business. 2019. Netflix adds 9 million paying subscribers, but stock falls. (2019). https://edition.cnn.com/2019/01/17/media/netflix-earnings-q4/index.html

[16] Francesco Cafaro. 2012. Using Embodied Allegories to Design Gesture Suites for Human-data Interaction. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. ACM, New York, NY, USA, 560–563. DOI: http://dx.doi.org/10.1145/2370216.2370309

[17] Children & Young People Now. 2011. Postcode lottery for care leavers setting up home. (2011). https://www.cypnow.co.uk/Other/article/postcode-lottery-for-care-leavers-setting-up-home

[18] Sandy Claes and Andrew Vande Moere. 2013. Street Infographics: Raising Awareness of Local Issues Through a Situated Urban Visualization. In *Proceedings of the 2Nd ACM International Symposium on Pervasive Displays (PerDis '13)*. ACM, New York, NY, USA, 133–138. DOI: http://dx.doi.org/10.1145/2491568.2491597

[19] Darius Coelho, Ayush Kumar, and Klaus Mueller. 2015. Data memes for personal visualization. In *Proceeding of IEEE VIS 2015 workshop Personal Visualization: Exploring Data in Everyday Life*.

[20] Andy Crabtree and Richard Mortier. 2015. Human Data Interaction: Historical Lessons from Social Studies and CSCW. In *ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work*. Springer International Publishing, 3–21. DOI: http://dx.doi.org/10.1007/978-3-319-20499-4_1

[21] Pedro Cruz and Penousal Machado. 2016. Pulsing Blood Vessels: A Figurative Approach to Traffic Visualization. *IEEE Computer Graphics and Applications* 36, 2 (Mar 2016), 16–21. DOI: http://dx.doi.org/10.1109/MCG.2016.29

[22] Ana C Andrés del Valle and Agata Opalach. 2005. The Persuasive Mirror: computerized persuasion for healthy living. In *Proceedings of the 11th International Conference on Human-Computer Interaction*.

[23] Chris Elsden, Bettina Nissen, Andrew Garbett, David Chatting, David Kirk, and John Vines. 2016. Metadating: Exploring the Romance and Future of Personal Data. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 685–698. DOI: http://dx.doi.org/10.1145/2858036.2858173

[24] Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean Munson. 2014. Taming Data Complexity in Lifelogs: Exploring Visual Cuts of Personal Informatics Data. In *Proceedings of the 2014 Conference on Designing Interactive Systems (DIS '14)*. ACM, New York, NY, USA, 667–676. DOI: http://dx.doi.org/10.1145/2598510.2598558

[25] Mike Evans, Tristan Ferne, Zillah Watson, Frank Melchior, Matt Brooks, Phil Stenton, Ian Forrester, and Chris Baume. 2017. Creating Object-Based Experiences in the Real World. *SMPTE Motion Imaging Journal* 126, 6 (Aug 2017), 1–7. DOI: http://dx.doi.org/10.5594/JMI.2017.2709859

[26] Brett Gaylor. 2015. Do Not Track: a personalized documentary series about privacy and the web economy. (2015). https://donottrack-doc.com

[27] Nahum Gershon and Ward Page. 2001. What Storytelling Can Do for Information Visualization. *Commun. ACM* 44, 8 (Aug. 2001), 31–37. DOI: http://dx.doi.org/10.1145/381641.381653

[28] Matt Golding. 2015. The Risk Taker's Survival Guide. (2015). http://risktakersguide.com

[29] David Philip Green, Simon Bowen, Jonathan Hook, and Peter Wright. 2017. Enabling Polyvocality in Interactive Documentaries Through "Structural Participation". In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 6317–6329. DOI: http://dx.doi.org/10.1145/3025453.3025606

[30] Jeffery Heer and Danah Boyd. 2005. Vizster: visualizing online social networks. In *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.* 32–39. DOI:http://dx.doi.org/10.1109/INFVIS.2005.1532126

[31] Nathalie Henry Riche, Christophe Hurter, Nicholas Diakopoulos, and Sheelagh Carpendale. 2018. *Data-driven storytelling*. CRC Press.

[32] Jonathan Hook. 2018. Facts, Interactivity and Videotape: Exploring the Design Space of Data in Interactive Video Storytelling. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '18)*. ACM, New York, NY, USA, 43–55. DOI: http://dx.doi.org/10.1145/3210825.3210826

[33] Dandan Huang, Melanie Tory, Bon Adriel Aseniero, Lyn Bartram, Scott Bateman, Sheelagh Carpendale, Anthony Tang, and Robert Woodbury. 2015. Personal Visualization and Personal Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics* 21, 3 (March 2015), 420–433. DOI: http://dx.doi.org/10.1109/TVCG.2014.2359887

[34] Noor Huijboom and Tijs Van den Broek. 2011. Open data: an international comparison of strategies. *European journal of ePractice* 12, 1 (2011), 4–16.

[35] Jessica Hullman and Nick Diakopoulos. 2011. Visualization Rhetoric: Framing Effects in Narrative Visualization. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (Dec 2011), 2231–2240. DOI:http://dx.doi.org/10.1109/TVCG.2011.255

[36] Marijn Janssen, Yannis Charalabidis, and Anneke Zuiderwijk. 2012. Benefits, adoption barriers and myths of open data and open government. *Information systems management* 29, 4 (2012), 258–268. DOI: http://dx.doi.org/10.1080/10580530.2012.716740

[37] Vaiva Kalnikaite, Abigail Sellen, Steve Whittaker, and David Kirk. 2010. Now Let Me See Where i Was: Understanding How Lifelogs Mediate Memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI âĂŹ10)*. Association for Computing Machinery, New York, NY, USA, 2045–2054. DOI: http://dx.doi.org/10.1145/1753326.1753638

[38] Helen Kennedy and Rosemary Lucy Hill. 2018. The Feeling of Numbers: Emotions in Everyday Engagements with Data and Their Visualisation. *Sociology* 52, 4 (2018), 830–848. DOI: http://dx.doi.org/10.1177/0038038516674675

[39] Christopher A. Le Dantec, Mariam Asad, Aditi Misra, and Kari E. Watkins. 2015. Planning with Crowdsourced Data: Rhetoric and Representation in Transportation Planning. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM, New York, NY, USA, 1717–1727. DOI: http://dx.doi.org/10.1145/2675133.2675212

[40] Ewen Macaskill and Gabriel Dance. 2013. NSA Files Decoded: What the Revelations Mean for You. (2013). http://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded

[41] Derek McAuley, Richard Mortier, and James Goulding. 2011. The Dataware manifesto. In *2011 Third International Conference on Communication Systems and Networks (COMSNETS 2011)*. 1–6. DOI: http://dx.doi.org/10.1109/COMSNETS.2011.5716491

[42] Sebastian Meier and Katrin Glinka. 2017. The Individual in the Data - the Aspect of Personal Relevance in Designing Casual Data Visualisations. *i-com* 16, 3 (2017), 247–265. DOI: http://dx.doi.org/10.1515/icom-2017-0025

[43] Sehl Mellouli, Luis F Luna-Reyes, and Jing Zhang. 2014. Smart government, citizen participation and open data. *Information Polity* 19, 1, 2 (2014), 1–4. DOI: http://dx.doi.org/10.3233/IP-140334

[44] Richard Mortier, Hamed Haddadi, Tristan Henderson, Derek McAuley, and Jon Crowcroft. 2014. Human-data interaction: The human face of the data-driven society. *Available at SSRN 2508051* (2014). DOI: http://dx.doi.org/10.2139/ssrn.2508051

[45] Bettina Nissen, John Bowers, Peter Wright, Jonathan Hook, and Christopher Newell. 2014. Volvelles, Domes and Wristbands: Embedding Digital Fabrication Within a Visitor's Trajectory of Engagement. In *Proceedings of the 2014 Conference on Designing Interactive Systems (DIS '14)*. ACM, New York, NY, USA, 825–834. DOI: http://dx.doi.org/10.1145/2598510.2598524

[46] Open Data for Development. 2016. International Open Data Roadmap. (2016). https://www.od4d.net/files/report-iodc-2016-web.pdf

[47] Open Data Institute. 2015. Open data roadmap for the UK. (2015). https://theodi.org/article/open-data-roadmap-for-the-uk-2015

[48] Andrea Polli. 2006. Heat and the Heartbeat of the City: Sonifying Data Describing Climate Change. *Leonardo Music Journal* 16 (2006), 44–45. DOI: http://dx.doi.org/10.1162/lmj.2006.16.44

[49] Zachary Pousman, John T. Stasko, and Micheal Mateas. 2007. Casual Information Visualization: Depictions of Data in Everyday Life. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (Nov 2007), 1145–1152. DOI: http://dx.doi.org/10.1109/TVCG.2007.70541

[50] Johnny Rodgers and Lyn Bartram. 2011. Exploring Ambient and Artistic Visualization for Residential Energy Use Feedback. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (Dec 2011), 2489–2497. DOI: http://dx.doi.org/10.1109/TVCG.2011.196

[51] María Teresa Rodríguez, Sérgio Nunes, and Tiago Devezas. 2015. Telling Stories with Data Visualization. In *Proceedings of the 2015 Workshop on Narrative & Hypertext (NHT '15)*. ACM, New York, NY, USA, 7–11. DOI:http://dx.doi.org/10.1145/2804565.2804567

[52] Edward Segel and Jeffrey Heer. 2010. Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics* 16, 6 (2010), 1139–1148. DOI: http://dx.doi.org/10.1109/TVCG.2010.179

[53] Tobias Skog, Sara Ljungblad, and Lars Erik Holmquist. 2003. Between aesthetics and utility: designing ambient information visualizations. In *IEEE Symposium on Information Visualization 2003 (IEEE Cat. No.03TH8714)*. 233–240. DOI: http://dx.doi.org/10.1109/INFVIS.2003.1249031

[54] Davy Smith, Jonathan Hook, and Marian Ursu. 2018. Authoring Object-Based Video Narratives. In *Adjunct Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '18)*.

[55] Thomas Smith, Simon J. Bowen, Bettina Nissen, Jonathan Hook, Arno Verhoeven, John Bowers, Peter Wright, and Patrick Olivier. 2015. Exploring Gesture Sonification to Support Reflective Craft Practice. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 67–76. DOI: http://dx.doi.org/10.1145/2702123.2702497

[56] John Stasko, Todd Miller, Zachary Pousman, Christopher Plaue, and Osman Ullah. 2004. Personalized Peripheral Information Awareness Through Information Art. In *UbiComp 2004: Ubiquitous Computing*, Nigel Davies, Elizabeth D. Mynatt, and Itiro Siio (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 18–35. DOI:http://dx.doi.org/10.1007/978-3-540-30119-6_2

[57] Alex S. Taylor, Siân Lindley, Tim Regan, David Sweeney, Vasillis Vlachokyriakos, Lillie Grainger, and Jessica Lingel. 2015. Data-in-Place: Thinking Through the Relations Between Data and Community. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2863–2872. DOI: http://dx.doi.org/10.1145/2702123.2702558

[58] The Care Leavers' Association. 2019. What is a Care Leaver? (2019). http://www.careleavers.com/who-we-are/what-is-a-care-leaver-2/

[59] The Care Leavers' Foundation and Children & Young People Now. 2009. Setting Up: A Place to Call Home. (2009). https://secure.toolkitfiles.co.uk/clients/23786/sitedata/files/Setting_Up_A_Place_to_Call.pdf

[60] The Children's Society. 2017. Give Care Leavers a Fairer Start: Campaign Resource Pack. (2017). https://www.childrenssociety.org.uk/sites/default/files/fairer-start-for-care-leavers-resource-pack.pdf

[61] The Children's Society. 2018. A Fairer Start for Care Leavers. (2018). https://www.childrenssociety.org.uk/what-you-can-do/campaign-for-change/a-fairer-start-for-care-leavers

[62] Lena Thiele. 2015. Netwars: A fact-based 5 part interactive web series exploring the impending threat of cyberwarfare. (2015).
http://www.netwars-project.com/webdoc

[63] Alice Thudt, Dominikus Baur, Samual Huron, and Sheelagh Carpendale. 2016. Visual Mementos: Reflecting Memories with Personal Data. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (Jan 2016), 369–378. DOI: http://dx.doi.org/10.1109/TVCG.2015.2467831

[64] Alice Thudt, Jagoda Walny, Theresia Gschwandtner, Jason Dykes, and John Stasko. 2018. Exploration and Explanation in Data-Driven Storytelling. In *Data-Driven Storytelling*. AK Peters/CRC Press, 77–102.

[65] Fernanda B. Viegas, Ethan Perry, Ethan Howe, and Judith Donath. 2004. Artifacts of the Presence Era: Using Information Visualization to Create an Evocative Souvenir. In *IEEE Symposium on Information Visualization*. 105–111. DOI: http://dx.doi.org/10.1109/INFVIS.2004.8

[66] Fernanda B. Viégas and Martin Wattenberg. 2007. Artistic Data Visualization: Beyond Visual Analytics. In *Proceedings of the 2Nd International Conference on Online Communities and Social Computing (OCSC'07)*. Springer-Verlag, Berlin, Heidelberg, 182–191. http://dl.acm.org/citation.cfm?id=1784297.1784319

[67] Annika Wolff, Matthew Barker, and Marian Petre. 2017. Creating a Datascape: A Game to Support Communities in Using Open Data. In *Proceedings of the 8th International Conference on Communities and Technologies (C&T '17)*. ACM, New York, NY, USA, 135–138. DOI: http://dx.doi.org/10.1145/3083671.3083686

[68] Rebecca Xiong and Judith Donath. 1999. PeopleGarden: Creating Data Portraits for Users. In *Proceedings of the 12th Annual ACM Symposium on User Interface Software and Technology (UIST '99)*. ACM, New York, NY, USA, 37–44. DOI: http://dx.doi.org/10.1145/320719.322581

[69] John Yorke. 2014. *Into the Woods: A Five-Act Journey into Story*. Abrams.

[70] YouTube. 2019. YouTube By The Numbers. (2019).

[71] Jack Zhao and Andrew Vande Moere. 2008. Embodiment in Data Sculpture: A Model of the Physical Visualization of Information. In *Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts (DIMEA '08)*. ACM, New York, NY, USA, 343–350. DOI: http://dx.doi.org/10.1145/1413634.1413696