

# More Weather Tomorrow. Engaging Families with Data through a Personalised Weather Forecast

Catho Van Den Bosch  
catho.vandenbosch@vrt.be  
VRT Innovation Belgium  
Belgium

Nikki Peeters  
nikki.peeters@vrt.be  
Flemish Public Broadcaster VRT  
Belgium

Sandy Claes  
sandy.claes@luca-arts.be  
Mediated Environments  
LUCA School of Arts Belgium  
Belgium



**Figure 1:** Close-up of application on touchscreen (left), father with two children interact (middle, left), touchscreen set up at the cafeteria (middle, right) and two children interacting (right).

## ABSTRACT

As data literacy skills are increasingly important in today's society, scholars have been exploring strategies to engage people with data, for example through storytelling and familiar media such as video. In this paper, we present the design of a video-based data storytelling application that prompts children and their families to explore and interpret historical weather data through a personalised weather forecast. The application was displayed at a 2-month summer exhibition of a popular television channel. In a controlled comparative study, we investigated how the application triggered reflection, as well as emotional and narrative engagement of families at home and at the exhibition. We combined this approach with an in-the-wild study, in which we observed spontaneous interactions of visitors. Our findings indicate that data engagement is encouraged when family interactions occur, which may be facilitated by external environmental conditions and internal story design. Here, we uncover 5 design recommendations for data video storytellers.

## CCS CONCEPTS

• **Human-centered computing** → **Interaction design theory, concepts and paradigms.**

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

IMX '22, June 22–24, 2022, Aveiro, JB, Portugal  
© 2022 Association for Computing Machinery.  
ACM ISBN 978-1-4503-9212-9/22/06...\$15.00  
<https://doi.org/10.1145/3505284.3529972>

## KEYWORDS

storytelling, visualisation, television/video, empirical study that tells us about how people use a story

## ACM Reference Format:

Catho Van Den Bosch, Nikki Peeters, and Sandy Claes . 2022. More Weather Tomorrow. Engaging Families with Data through a Personalised Weather Forecast. In *ACM International Conference on Interactive Media Experiences (IMX '22)*, June 22–24, 2022, Aveiro, JB, Portugal. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3505284.3529972>

## 1 INTRODUCTION

Today's greatest challenges, such as global warming and the Covid-19 pandemic, cannot be fully grasped in everyday life as the significance of the related data remains abstract to a certain degree [1]. To understand what the presented data mean, citizens need additional support and guidance [23]. The ability to read, interpret and understand data related to different topics is therefore considered key to participate in an increasingly data-driven world [38] as it allows citizens to comprehend governmental issues, participate in civic debate [10], and take informed decisions. Furthermore, it strengthens their inclusion as well as resilience in future society. The pandemic has shown us that datafication is accelerating, and that such data literacy skills are more urgently needed - especially in a media context [22]. Among others, data representations are used by different media to present policy impact, and are deemed crucial for instructing people's behaviour [30]. Studies have shown that infographics, such as the "flatten the curve" example used during the Covid-19 pandemic, can cause non-expert audiences to elaborate more than when using purely textual and visual messages

[17]. However, audiences are not as data literate as media producers expect them to be, causing misconceptions that might lead to a decrease of trust [22] or the spread of misinformation [9].

Data-driven storytelling techniques offer a reliable and engaging approach to enhance data literacy skills [4, 12, 29], but media organisations are not widely adopting these approaches. It requires a too progressive shift within the organisation in terms of content, production tools and presentation platforms [34]. As journalists and documentary makers are accustomed to aggregating data into video stories, the inclusion of interactive data representations in video offers a more accessible way for media producers to transition to data storytelling [15]. Another advantage of video is the presence of story characters that viewers can relate to, which makes them empathise with the shared thoughts, and triggers critical reflections on those experiences [21]. These feelings of empathy and consideration can be stimulated even more when you personalise stories based on the viewer's situation. [8]. Personalised data videos thus offer an entry point to engage both media producer and viewer with data. Despite this, their design and particularly the relation to their context are still relatively under-explored [8]. For example, in documentary film and journalism, storytelling entails more than the story as an object, it also includes the elicitation and construction of the objects in situ [21]. This construction (including the decision to engage or not) depends on the context, such as the audience, purpose, and location. Storytelling emphasises the performance it entails, such as a social experience that involves retelling with slightly different interpretations [21]. However, few studies have explored how different settings, such as a screen at home or a physical installation at an exhibition; and related social aspects, such as interacting with children, family or friends, facilitate engagement with personalised data stories, and in particular, how this affects critical reflection.

In this paper, we present the design and evaluation of a personalised data video story in the form of a weather forecast that invites children and their families to interpret historical temperature and precipitation data. As we encounter weather data every day in various formats (e.g. weather apps, news bulletins, etc.), we chose to work with this familiar data set in order to address a broad audience [19]. More specifically, the design of our personalised data video focused on both children and adults in order to encourage families to experience the data video together, and thus collectively make sense of it. We describe the implementation of an application that facilitates a personalised weather forecast on a touchscreen, which was showcased at a 2-month temporary interactive exhibition on weather, organised by a family television channel. We executed a comparative study with the same application at home, for which we each invited 19 (groups of) participants. Next to this, we conducted an in-the-wild study at the exhibition for 5 days. We sought to explore how the external context and internal story design influenced their narrative engagement, emotional commitment and contemplation. By looking at our findings through the lens of traditional video storytelling techniques, we present 5 design recommendations to connect the design of data representations for open exploration to the performing character of video storytelling. Challenges such as global warming and the Covid-19 pandemic need to be addressed collectively; social dynamics are thus invaluable for experiencing, interpreting and sharing data stories. Therefore, we believe our

findings contribute to the understanding of how data videos elicit collective consumption and sense-making of data.

## 2 RELATED WORK

Data-driven storytelling techniques are already well-studied to inform the design of data visualisations for a broad audience [29]. These studies focus on data stories that are distributed via online platforms or newspapers, while other forms of popular media, such as television, are often overlooked - despite their potential to reach a highly diverse audience [11]. Scholars have been exploring the role of video in data stories [2, 15], and more particular novel interactive distribution means, such as responsive video technology [8], which is also being studied in television contexts (e.g. [37]). One of the benefits of data videos is their ability to emotionally engage viewers with characters that act as participants in the data story, e.g. [8] or as presenters, e.g. [15], which triggers reflection [8] and stimulates data exploration [5].

Next, (interactive) television and data videos are often consumed together with others. In data storytelling, scholars recently explored how design can further stimulate collaborative analysis, data reflection and social discussion [25]. Others explored the specific setting of classrooms as a way to collectively analyse data through personalisation [36] or layered story design [18]. In television and film, the genre of family film often takes the form of animated film or video, and has mastered the ability to connect with children and adults simultaneously [20]. Before video streaming platforms were omnipresent, television channels, and especially public service media channels, directed the programming of their Saturday evening to families, with the goal of establishing social interaction between them through game shows and comedy. Similarly, museums are also involved in designing narratives for family experiences [41], which might be exploited in data exploration contexts as well [35].

## 3 MORE WEATHER TOMORROW

In general, we are accustomed to encounter weather data; we consume weather forecasts and consult weather data applications on a daily basis. [19]. Moreover, weather forecasts are often consulted in a family context, e.g. when to decide what to wear or what to bring along for activities, and are also part of everyday conversations with familiar strangers [16]. Here, current meteorological data, including temperature, wind direction and speed, amount and type of precipitation and hours of sunshine [14] are interpreted by weather presenters or journalists, who add meaning to the data and present their interpretations in an informative story.

The goal of our personalised data video installation was to engage families in interpreting weather data together, and reflect upon the meaning of their interpretation without a particular goal in mind, such as adopting a certain mindset towards climate change. We deployed a storytelling approach to guide viewers through this open exploration, which we will elaborate on in the following part. We will separately discuss (1) external story design factors that storytellers take into account, such as location and time of day of story exposure, and (2) internal story design elements, such as the deployment of three acts. Yet note these factors are dependent on each other and cannot be seen separately.

### 3.1 External story design

External elements that impact story design include context [21], which is provided in this study by the family exhibition 'More weather tomorrow' (translated from the Dutch phrase 'Meer Weer'). The title of the exhibition is based on the catchphrase from a weather presenter of the family television channel *Eén* in Belgium. The exhibition aims to inform families about different weather phenomena, such as durability and climate change, through playful didactic installations. It ran for 2 months during the children's summer break and was located at a popular summer venue.

The physical installation that consisted of a 43-inch touchscreen on a pedestal, was presented at the end of the exhibition, in the cafeteria. Overall, the exhibition covers 10 interactive installations and a trajectory of 40 large-scaled photo backdrops that allowed visitors to take spectacular photos of weather-related phenomena (e.g. in a thunderstorm but also outside a spaceship) as a way to encourage them to read facts about these phenomena.

In the cafeteria, 4 installations competed for attention with our installation, as well as other surrounding stimuli such as a bar area, people and non-digital signage.

The producers were aware that the application was designed and developed as a *research product* with the aim to gather real-world insights [24]. Apart from our application, the youth news and weather forecast was not represented at the exhibition. Instead, the regular, adult-oriented weather forecast of the main family television station was visually apparent. The producers were convinced that this application could deliver an added value both in terms of educational goals as in the visual presence of the youth news, including the weather forecast. Therefore, we were able to collaborate with the youth news journalists who provided feedback on our ideas, and aided the overall story design and deployment of storytelling techniques, which we will discuss in the next section.

### 3.2 Internal story design

The story that was shown on the display (i.e. internal) followed a traditional three-act structure, i.e. begin, middle and end. We made sure the data video was not heavily author-driven, as this could diminish the qualities of data representations as a mirror of the objective truth [8]. Instead, it should facilitate different reader-driven possibilities to think or reflect upon the data by interacting with it [31].

**3.2.1 The Beginning - a Personalised Infographic Animation. Setting the scene.** The application uses the same design language of the existing (linear) youth news weather forecast that is aired every day on the children's channel via broadcast television, their streaming platform and their website. The square shapes, bold colours and dynamic motion design are typical for this brand, and are therefore recognisable for a large audience. The graphic design exposes the context in which the story will take place, i.e. the weather forecast of the youth news, and sets expectations for the events that will follow.

**The Hook.** The application asks to "discover what the weather was like on your date of birth", complemented with 3 large squares with which people can manipulate the showcased year, day and month by clicking the arrows or selecting the date numbers via a calendar pop-up. Then, it calculates the weather forecast for

that exact day, which is made visual through a metaphorical slot-machine animation, followed by the display of the result in 3 pink squares (see Fig. 2b). This result is presented as an infographic via a series of icons that depict the overall weather from sunny, cloudy, raining a little, raining a lot or snowing, the temperature in a number of degrees Celsius, to the items you would need that day, including an umbrella, a raincoat, a t-shirt, a sweater or a coat. The direction of wind is also shown below the squares. The infographic is identical to that of the daily youth news weather forecast, which we consider as a comprehensible entry point as it is similar to other weather applications or newspaper graphs.

**3.2.2 The middle - Personalised Videos and Line Graphs. Zooming out** A weather presenter often explains the underlying reasons for the past, current and future conditions of the weather and their individual data points, by zooming out to present the bigger picture [21]. This part of the story is split into a second and third act (since the beginning is the first act), in which the complexity of the data representations increases.

**Second act** After the infographic result is shown on the application, a button pops up in the right corner, allowing the visitor to learn more. An appropriate video is loaded based on the outcome of the birth date. Here, the presenter of the youth news, i.e. Jelle welcomes the viewer with its familiar presence, increasing interaction and trust [13]. Then, Jelle says, for example: "When you were born, people were wearing raincoats. Shall we take a look if this has always been the case on that day?", after which a line graph unfolds with historical precipitation data from the respective birth year up to 2020.

The line graph is more complex to interpret than the infographic in the beginning, particularly for children. Yet, at the age of 12, children can have the same interpretation skills as the lay adult audience, while children between 9-11 years old are able to understand graphs that are not bound to judgements on volume [32]. A line graph is still complex for children as it requires them to understand concepts such as 'larger than'. On the other hand, it looks familiar, which is also a quality that supports a broad audience to engage with data representations [27]. The line graph is complemented with three buttons: (1) a query button that indicates "Do you want to compare with Jelle's birth date?", which reveals a second line based on the evolution of weather on the 20th of April, and (2) a stepper button "What with the days before you were born?" that progresses the line graph to the start count of the data set in 1920. When this larger period is shown, an additional query button allows the visitor to activate relevant milestones in history, such as the invention of television, in order to add context to the time spanning the data set. A third (3) button allows the viewer to continue to temperature data.

**Third act** Jelle takes over the screen in video. Once more, depending on the outcome of the weather forecast of your birth date, he will present a particular insight that you could have gained from the chart, e.g. "It was raining more on the day of your birth than on the same day in the subsequent years. But what about the temperature?" The application continues to show a line graph that presents the historical temperature data since your birth year. Then, the same order of query and stepper buttons as in the previous 'act' are shown with an additional query button in the end "How did the



**Figure 2: Entry form for birth date (a), weather forecast for that date (b), the weather presenter in a video still (c) and evolution of precipitation data for that date (d). Titles are translated from local language to English and included in a different typeface than the original application.**

*weather evolve*", which reveals a trendline. The latter is even more complex to interpret as most dates only show a slow ascendant line.

**3.2.3 The End - Personalised Concluding Video and Call-to-Action. Resolution** Finally, the forecast ends with a concluding insight of Jelle, depending on the historical temperature data. For example, he says *"The temperature at your birth date has increased over the years. That is caused by the climate change, about which you can learn more at this exhibition"*.

As our application is based on the birth date of the viewer, we needed to prepare for different 'plot twists'. Therefore, we recorded 10 different videos with Jelle; a practice that was novel to the youth news producers. These videos were edited in Adobe Premiere Pro and After Effects. Depending on the outcome, one of 5 videos in WebM (or .mp4, depending on supporting browser) was dynamically loaded. The line graph that followed was implemented with D3.js. Other features such as the stepper and query button, and the unfolding of the graphs were implemented using HTML, JavaScript and JQuery. Titles and labels were also made available in a second local language, which viewers could select in the beginning of the application (before entering their birth date). As a result, the video of the presenter would also be subtitled (see Fig 2. c).

The historical data that were fed into the application was retrieved from an open source data platform; a crucial aspect for our design goals as people need to be able to find the source data. However, public service media have their own data licenses and typically deploy data that are not openly accessible. We thus needed to defend our choice for open data, which demonstrates that to truly incorporate data representations that encourage open exploration in mass media channels, agreements with existing data delivery services also need to be revisited.

## 4 METHODOLOGY

We conducted two studies, following a mixed-method approach: 1) a comparative study of the experience at the exhibition versus at home, and 2) an in-the-wild study at the exhibition. Ahead of these studies, we evaluated the performance of the personalised weather forecast in a pilot study to identify potential issues with story flow and data interpretation. Semi-structured interviews with 8 participants (all younger than 23 y.o., with 4 participants younger than 14 y.o.) allowed us to refine the graphic design. This also

motivated us to add 2 extra points to the graph, i.e. the wettest and warmest birthday over the past 100 years, to allow viewers to immediately spot outliers.

### 4.1 Comparative Study

We executed a comparative, between-subject design study with 19 participants per condition, i.e. at the exhibition (CE) or at home (CH). We adapted the application to be experienced online, via the web browser, and changed the story accordingly (i.e. no reference to the exhibition was made). We were aware the interaction modality shifted from only touchscreen to either touchscreen (on a tablet) or a mouse and keyboard (on a computer). As we aimed to relate to realistic settings at home, both in terms of distribution (i.e. via online platforms) and consumption (i.e. via personal and varying devices), we optimised our application for this experience. The study procedure was identical in both conditions. After consuming the data video story, participants completed a 7-point Likert questionnaire that included engagement measures commonly used in the domain of film and television to reflect on the role of storytelling on emotion, narrative engagement and contemplation [3].

**Contemplativeness scale.** The first 9 questions (see Fig.3, top) were intended to measure the level of reflective thinking triggered by the application through a contemplativeness scale. [3]. Questions included the level of involvement and the amount of new insights gained by the users.

**Emotional engagement scale.** The next 6 questions aimed to gain insights in the emotional engagement with the characters, i.e. the presenter of the youth news [3]. This was further explored through discussions with other participants (see Fig.3, middle).

**Narrative engagement scale.** The last set of 8 questions (see Fig.3, below) mapped the subjective experience of participants through the narrative engagement scale [6]. This scale aims to examine how external distractions affected the amount of focus the user had and how the narrative structure influenced the flow.

**Open questions and demographics** We asked the participants 4 open questions in which they could formulate their discoveries (*"What new information did you discover?"*, *"What did you remember?"*, *"How confident are you with interpreting data representations in popular media?"*, *"What did you already know about this subject?"*). Finally we asked their demographic data, such as their age and their relation to those that accompanied them.

Two hypotheses guided this study:

*H1. Participants report higher levels of narrative engagement at home than those at the exhibition.* We expect that the busy environment of the cafeteria, i.e. where other visitors are walking and talking out loud, would distract viewers from consuming the full data story. We also expect that more social interaction and collective data interpretation with family members would occur at the exhibition than at home, where they are also more distracted from narrative construction, similar to studies in sport video stories [42].

*H2. When emotional engagement of participants is high, social reflection is triggered.* We assume that when viewers engage with a data story together with family or friends, they reflect on the relevance of the presented data in relation to their family, community or society as a whole, and discuss it with them. Furthermore, we expect viewers to be more emotionally engaged with the topic at the exhibition, since it immerses visitors in a specific atmosphere.

For the condition at the exhibition (CE), we recruited 19 adult participants at the entrance of the exhibition. The sample consisted of an adult with at least one child younger than 10 years ( $N = 10$ ), at least one child between 10 and 16 years old ( $N = 4$ ), with children older than 16 ( $N = 4$ ), or alone ( $N = 1$ ). We asked them to test the full application after which they were given a tablet computer to fill in the post-survey. On average, this participation took 4 - 6 minutes of time, for which they were rewarded with a drink coupon at the cafeteria.

For the condition at home (CH), we invited viewers of the Flemish youth channel *Ketnet* via the broadcaster's recruiting platform, i.e. *VRT Pilotzone*. After the participants at home tested the application they were asked to complete the post-survey. On average, filling in these questionnaire took 3 - 4 minutes. Participants in this condition were not rewarded. We explicitly asked the 19 participants to test the application together with family members, including at least 1 child younger than 10 years ( $N = 6$ ), children older than 10 years ( $N = 4$ ), or with people older than 16 years ( $N = 1$ ), yet some tested it alone ( $N = 6$ ). There were also participants of the study who did not answer this question ( $N = 2$ ). We wanted to ensure that the application was tested with a child, so our recruiting efforts were focused on the associated target audience through the youth channel and its branding. Participants also first confirmed their ability to test it with a child (as we could not invite children directly due to ethical issues). Despite this, at the end of the test, 6 participants reported to have tested it alone. Potentially, the children dropped out during the testing; a situation that also happened at the exhibition.

## 4.2 In-The-Wild Study

The aim of this study was to gain insight in the spontaneous behavior of groups of people that engaged with the application, observe the social interaction around the physical installation and gather honeypot effects [40]. We organised participant observations during 5 days (when the exhibition opened at 11am till closing time 6pm) randomly spread over the last 4 weeks of the 2-month set-up, including one Tuesday, Thursday, Saturday and two Fridays. On those 5 days, 2 researchers were alternating their presence in the cafeteria to manually log engagements with the application. The observer was seated at a table of the cafeteria, 3-4 meters away

from the installation, and wrote down the number of visitors that halted, whether they engaged and for how long, their behavior, and their conversations. There was a sign that informed visitors of the installation of possible observations taking place.

The actions and conversations of 30 groups and 8 individuals were described in detail. The groups were mainly parents and children ( $N = 15$ ) and parents and grandchildren ( $N = 9$ ). There were also groups of children ( $N = 3$ ) and other compositions ( $N = 3$ ). The researcher present approached 10 of those groups that had spontaneously engaged with the application, to report on what they had discovered, in order to collect their open and spontaneous responses. On average, these interviews took 3-4 minutes.

## 4.3 Analysis

**4.3.1 Quantitative results.** We conducted a Cronbach's alpha test that revealed both scales are consistent (alpha home = 0,81, alpha exhibition = 0,98). For all those questions the mean and standard deviation for the two conditions (CE = Condition Exhibition, CH = Condition at home) were calculated and compared. For specific questions, a Wilcoxon Rank-Sum test was used to determine if the difference was significant.

**4.3.2 Qualitative results.** The reported insights coming from the open questions in the post-survey were independently coded by two researchers, including the main author, according to the types of visualisation insights triggered by personal data [7]: 'detail', when referred to specific data points, 'trend', when reporting data changes over time and 'data summary', when summarising all data. 'Self-reflection' includes different subtypes, such as connecting data presented to uncaptured data, contradicting with existing knowledge or confirming existing knowledge. Also the type 'data comparison' includes several subtypes, such as comparing with other data, or bringing in external data for comparison. Then, the resulting codes were discussed by the main author and another researcher, until they reached an agreement. In the coding and the discussion, we observed how insight formulation differed between conditions. Therefore, our analysis included whether they were written in a distant and factual manner, personal or social (i.e. when data points of at least 2 persons were explicitly mentioned). The in-the-wild observation annotations were analysed independently by the main author and a second researcher (who was not present at the exhibition) via open coding [33], resulting in 21 thematic codes that were refined by two authors in 3 iterative sessions. Because of this, certain tendencies were discovered, which contributed to the results and discussion section. Overall, it should be noted that not all entries were considered insights. In both conditions, participants also made comments on the design, e.g. "*The subtitles of the presenter were not accurate*" or shared opinions about what they discovered "*It was interesting*", or, "*I did not learn anything new*". We would also like to point out that the analysis was performed in the local language, and then translated to English.

## 5 RESULTS

### 5.1 Quantitative results

Visitors of the exhibition typically spent 2 minutes and 39 seconds using the application (i.e. median of all times the application was



started during the 2 deployment months ( $N = 6,497$ ). The median time spent with the application in CH was 1 minute and 14 seconds. The median is chosen here instead of the mean, to exclude participants who closed the application immediately or not at all. In terms of their experience with interpreting line graphs, 5 participants in CH reported they have experience with interpreting line graphs, 2 answered that this was not common in their daily lives and 7 did not fill out this open question. In CE, 6 participants reported to have experience interpreting graphs while the others were not that familiar with charts and statistics.

The participants in both conditions indicated that they generally understood what the application was about (CE: Mean = 6,29, SD = 0,89, CH: Mean = 6,26, SD = 0,85) and they were mainly motivated to explore the application as they cared about the subject (CE: Mean = 6,18, SD = 0,92, CH Mean = 5,06, SD = 1,13). As shown in figure 3, all parameters of the contemplativeness scale scored higher for the exhibition condition. The participants who were present at the exhibition felt more involved (CE Mean = 5,88, SD = 1,08) than the participants who tested the application at home (CH Mean = 5,06, SD = 1,30). On average, CE participants also thought more about the societal challenges related to the topic than in CH (CE: Mean = 4,53, SD = 1,65, CH: Mean = 3,21, SD = 1,79).

At the exhibition, participants were inspired to talk about the evolution of weather data with others (such as family, friends,...) (CE Mean = 5,05, SD = 1,88), which was significantly less at home (CH Mean = 3,37, SD = 1,63) according to the Wilcoxon Rank-Sum Test ( $p=0,01078$ , significant  $p < 0,05$ ). In addition, they were also more encouraged to further explore the topic (CE Mean = 5,05, SD = 1,67), while this occurred less at home (CH Mean = 3,53, SD = 1,79).

The emotional engagement that was triggered by the presenter proved to be neutral in both conditions; participants hardly identified themselves with him (CE mean = 3,42, SD = 1,96; CH mean = 2,21, SD = 1,47). There was one exception: children under the age of 10, who participated under supervision of their parents, felt more connected to the presenter and his ideas than other participants (mean = 4,25, SD = 1,92).

Participants at home found it less logical and convincing (CH Mean = 4,16, SD = 1,81) than at the exhibition (CE Mean = 5,00, SD = 1,81). Participants of CH were also more distracted by their surroundings (CH Mean = 3,05, SD = 1,76), which was significantly different for the participants at the exhibition (Mean = 1,84, SD = 1,14), according to the Wilcoxon Signed-Rank Test ( $p=0,03318$ , significant  $p < 0,05$ ).

## 5.2 Qualitative results

As shown in Fig.4., CH participants were more focused on reporting insights as details (CH  $N = 8$  (out of 21); CE = 6 (out of 30)) e.g. *"It was cold on the day I was born"* - CH P12, while at the exhibition, more people talked about trends e.g. *"We saw that our birthdays got warmer and warmer over the years"* - CE P11 (CE  $N = 7$  (out of 30, CH  $N = 4$  (out of 21)). In terms of self-reflection, CE participants often made comparisons with the external context e.g. *"Our mom always told us that it was raining a lot when I was born, well, that is not true"* - CE P15 (CE  $N = 10$ ), while this did not occur with the insights formulated by the participants at home (CH  $N = 0$ ).

We observed 5 times when an individual or a part of a group started testing the application and then invited the rest of the group to join e.g. a child to a parent *"When were you born? We can have a look what temperature it was then"* - CE P8. The family experience was also mentioned by 7 CE visitors as a reason to sustain their engagement e.g. *"I found it very exciting. I am also very interested in the weather. And I try to pass that on to my son"* - CE P6.

CE visitors ( $N=7$ ) retold the data stories to family bystanders after they finished using the application, for example sharing with each other when the rest of the group was seated in the bar e.g. *"Look it was often cold when you had your birthday"* - CE P15, *"I saw that it was 20° on my birthday"* - CE P13.

The exhibition was targeted at a young audience, whose knowledge of graphs and data is sometimes limited. When asked what she thought of the graphs, a 10-year-old girl replied: *"I do not even know what graphs are"* - CE P2. We noted 4 occasions in which (grand)parents explained to their children the meaning of the different lines. One father mentioned: *"I explained them how the temperature changed of the years, the trend line is a good simplification to show that the temperature has gotten warmer and warmer"* - CE P16. In another occasion, a grandfather reported: *"I have worked in a school myself and I find this a good way to teach them how things have changed."* - CE P11. We also saw this dynamic when the digital skills of the grandparents were limited or when they were not that interested in news e.g. *"I think that those kids know more about it than us, they watch the news at school"* CE P2.

We interviewed 9 visitors who took photos of the personalised weather infographic for family or friends at home, after which they expressed to share it with others 'as proof' ( $N = 5$ ), or as an interesting fact for family members who were not present ( $N = 2$ ).

In CE, 10 interviewed visitors reported their discoveries in a story that departed from a hypothesis or previous experience, e.g. *"I expected this [...], contrasted with the data (and its evolution), e.g. "but then [...], followed by a conclusion "so it was true" or action, e.g. "I am going to send mom a photo of this"* - CE P16.

At CE, some ( $N = 8$ ) made hypotheses that were based on previous experiences or existing knowledge e.g. *"I think it will be cold on my birthday"* - CE P18 or on family stories e.g. *"I heard it was very hot the day Aunt (name) was born, it was apparently a heat wave then"* - CE P19. We noted in 5 occasions that these hypotheses were formed while queuing. In CH, 3 participants indicated they departed from a particular assumption or existing knowledge based on memories.

CE visitors also entered other episodic dates than birthdays, such as summer holidays, wedding dates and when loved ones died, which they then compared with their memory of the event.

## 6 DISCUSSION

Our findings reveal how external and internal story design elements support data engagement.

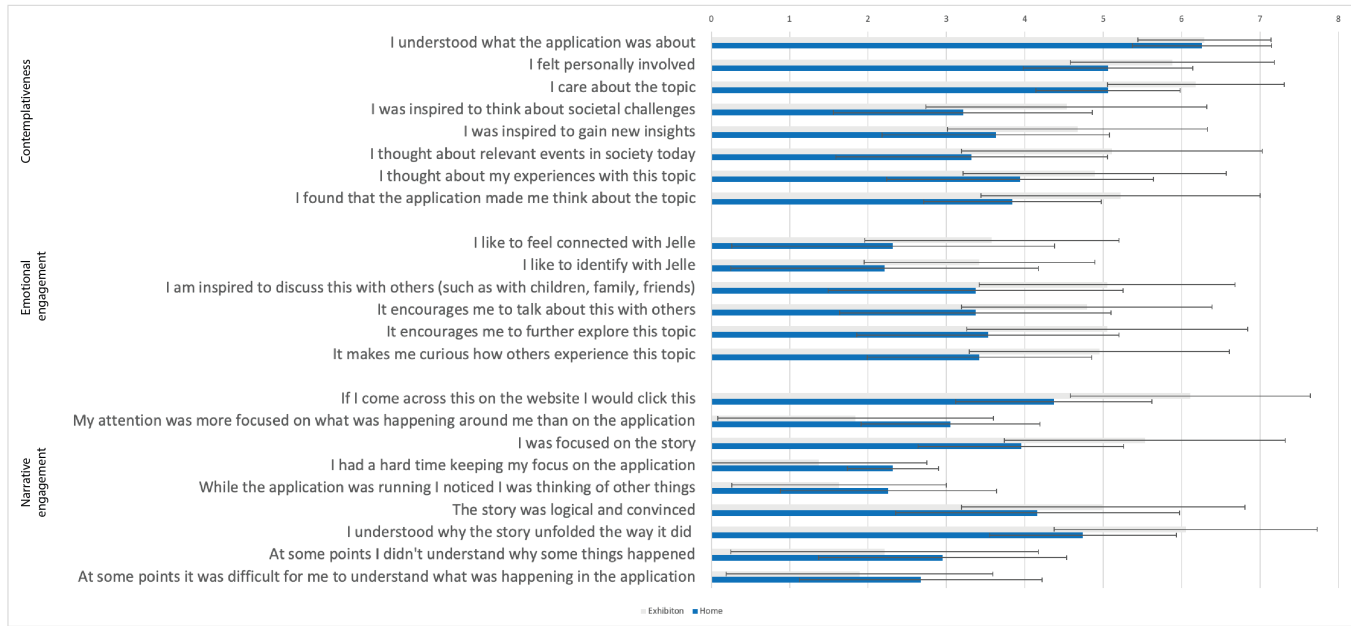


Figure 3: Experiences when testing the application

|  | Condition Exhibition | Condition Home |
|--|----------------------|----------------|
| <b>Detail</b>  | 6                    | 8              |
| e.g. "It was cold on the day I was born"   |                      |                |
| <b>Trend</b>   | 7                    | 4              |
| e.g. "We saw that our birthdays got warmer and warmer over the years"  |                      |                |
| <b>Self reflection</b>   |                      |                |
| Insights based on personal info other than birthdate   |                      |                |
| <b>External context</b>  | 10                   | 0              |
| e.g. "Our mom always told us that it was raining a lot when I was born, well that is not true"   |                      |                |
| <b>Contradiction</b>   | 1                    | 1              |
| e.g. "The temperature on the day I was born was not correct, it was much warmer than the application revealed"                         |                      |                |
| <b>Confirmation</b>  | 3                    | 0              |
| e.g. "It is true that it does not rain that much in August, actually the weather has always been nice on your birthday"                |                      |                |
| <b>Comparison</b>  |                      |                |
| <b>By factor</b>   | 1                    | 0              |
| e.g. "There was a clear difference between the forecast of our 2 daughters, but one is born in the summer and the other in the winter" |                      |                |
| <b>By time</b>   | 3                    | 2              |
| e.g. "The temperature at my birthdate stayed more or less the same, while the hottest day was already some years ago"                  |                      |                |
| <b>Against external data</b>   | 1                    | 3              |
| e.g. "My birthday is warmer because of climate change"   |                      |                |
| <b>Data summary</b>  | 2                    | 1              |
| e.g. "Climate change causes the temperature to increase"   |                      |                |

Figure 4: Different types of insights obtained by the participants

## 6.1 Unpredictability of the environment encourages families to engage longer with data

In contrast to our hypothesis (H1), participants were overall more engaged with the narrative at the exhibition, than at home. Beforehand, we expected that the loud atmosphere and the vicinity of several other interactive applications at the cafeteria would not keep their attention to the story, yet participants at home felt more distracted by their environment than at the exhibition. The time spent per condition also indicates that the external stimuli at home were more distracting. Moreover, the cafeteria environment allowed visitors to invite others, who were not involved from the start, to

join the experience. For family members who were already experienced with interpreting data, the unfamiliar and unpredictable surrounding provided an additional layer of challenge in the activity at hand and encouraged family interaction [20] and engagement with the data.

**Design recommendation 1.** Data video storytellers should consider external environmental conditions, such as time and location, as an additional narrative layer for those who are familiar with the main story line. Since television and media nowadays are mainly experienced on-demand, it is difficult for storytellers to predict how and when their story is experienced, in contrast to traditional linear stories that families may enjoy together on Saturday evening.

To overcome this, we imagine storytellers could include cues that encourage young viewers to engage their parents in the data experience, by requesting their participation and as such, trigger family interaction. We also believe that designing for family interaction would be more effective when it can be experienced at devices best suited for family consumption, e.g. a television screen on which the application can be experienced via a streaming platform. Future work could further explore the role of the distribution platform in data engagement and exploration, for example by creating interactive (video-based) data stories for television.

## 6.2 Connecting data to memories support reflection

Our insight analysis revealed insights in both conditions equally grasped the context of global warming. However, the contemplativeness scale per condition indicated that participants at the exhibition contemplated more about societal challenges related to the data topic than at home. We thus overestimated the role of social discussion in reflecting on the social relevancy of the topic. The exhibition environment did not contribute to emotional engagement with the topic. Instead, social discussions on existing factual knowledge, memories or family stories, often formed the starting point for participants' discoveries. As discussed in the results section, the hypothesis formulated by the participants in both conditions helped them to draw conclusions from the data. We therefore believe linking data to memories is a strong narrative device to trigger data reflection. Our memory is story-based, which can be triggered through touchpoints. In traditional video storytelling, such touchpoints are typically provided as a way to make the audience relate to the characters [39]. In this case, the data video story provided memory touchpoints that connected the viewers with the data, facilitating reflections in a personal or social manner.

**Design recommendation 2.** Memories form a compelling narrative device to encourage reflection. Data video storytellers should consider establishing a connection between the episodic memories of their intended audience and the data. Here, data should not only be considered as an end point, they are a medium to share personal or social experiences [26].

## 6.3 Connecting data to memories support retelling

Reported insights that are complemented with memories form fun data exploration points - whether they turn out to be true or not - to be shared with others who also may recall the occasions. Our findings revealed how visitors retold the data story from their own point of view to their family, and even to us, researchers. These stories would often start with a personal hypothesis based on a memory, followed by the actual data and a reflection on its meaning. Thus, the episodic memory formed the basis to retell the data story. Our findings also show that both participants and visitors were encouraged to enter other episodic dates, such as summer holidays, wedding dates and even dates when loved ones had died, which they then compared with the objective data. We found several visitors took photos of the personalised weather infographic of family or friends at home, and shared it via social

media. The episodic memory thus forms a good basis to retell the data story.

**Design recommendation 3.** Sharing practice encourages humans to retell the data story, thereby describing data points and insights. As a result, data story (re-)tellers are familiarised with data. Designers of data video stories could consider including elements that respond to memories or other social discussion points, such as famous persons, to facilitate the retelling of data stories.

## 6.4 A familiar presenter establishes trust and encourages data exploration

At the exhibition, we learned how the presenter's presence and call to explore the data became an additional nudge to interact with the line graphs. For (grand)parents, the inclusion of the familiar youth news presenter established trust, and made them encourage their children to further engage with the application. In addition, when participants recognised the presenter when others were using the application, they also became driven to find him themselves through the entire data video story. Including a familiar character, such as a presenter, in a data story formed an additional motivation to explore the design for those who relate to them (similar to the presence of famous actors in movies), which appealed to a subset of the audience. It forms an additional incentive to sustain data engagement. In addition, a familiar news presenter is a particular way to establish trust in the information that is conveyed, which is a strategy already used in the fight against disinformation [13].

**Design recommendation 4.** When aiming to engage children with data, including a familiar or popular character in the story design can form an additional entry point for data engagement.

## 6.5 Exchanging data-literacy skills supports family interaction

We observed moments of inter-generational learning by passing on data literacy skills at the exhibition. Such inter-generational learning also occurred from children to their (grand)parents. The integration of different levels of data visualisation complexity allowed a data-literate audience to connect to the data video story and be triggered to pass on their knowledge. Storytellers that write family films are already experienced in designing stories that connect both to children and adults through different narrative layers that meet the needs for knowledge of each type of viewer, i.e. adult or child [20]. This is in line with recent findings on integrating layered storytelling approaches for data analysis in collaborative learning contexts [18].

**Design recommendation 5.** Data video storytellers that aim to engage families should consider this connection to different data literacy levels, next to the more traditional narrative layers of the internal story design that were discussed in design recommendation 2, 3 and 4.

## 6.6 Limitations

Often in media design, effects are topic-related, causing them to be difficult to generalize [28]. However, our study demonstrated how personal and social connections to data, such as important event dates, memories and affection for particular characters, can form



powerful entry points for data engagement. These design recommendations are relevant for other topical data sets. For instance, we imagine graphs that depict Covid-19 cases related to personal (e.g. first ill person in the family) or social events (e.g. first party with limited people) are a way to draw people in and interpret data. Similar scenarios can be thought of for other news data, such as elections, financial news and sports data, to name a few.

## 7 CONCLUSION

In this paper, we present the design of an interactive weather forecast that reveals historical data of temperature and precipitation by means of personalised video storytelling techniques. Through a comparative study in a home and family exhibition environment, and an in-the-wild study at the exhibition, we learned that family interaction supports social reflection on the data presented, and that environmental conditions affect data engagement. Furthermore, our findings suggest that establishing a connection between data and memories is a compelling device to foster engagement with data, and support retelling of data stories, and that a trusted, familiar presenter can help overcome reticence with more complex types of data representations. These findings suggest that professional data storytellers do not only need to consider the narrative, internal story design but also the external conditions of story consumption.

## ACKNOWLEDGMENTS

We thank Quinten Mares (Innovation Department of VRT), Kyo Schelfhout (student Media and Information Design at LUCA School of Arts) for the design and development of the More Weather application and the people at VRT.

## REFERENCES

- [1] Karen Akerlof, Edward W Maibach, Dennis Fitzgerald, Andrew Y Cedeno, and Amanda Neuman. 2013. Do people “personally experience” global warming, and if so how, and does it matter? *Global environmental change* 23, 1 (2013), 81–91.
- [2] 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* (Seoul, Republic of Korea) (CHI '15). Association for Computing Machinery, New York, NY, USA, 1459–1468. <https://doi.org/10.1145/2702123.2702431>
- [3] Anne Bartsch, Anja Kalch, and Mary Beth Oliver. 2014. Moved to Think. *Journal of Media Psychology: Theories, Methods, and Applications* 26 (01 2014), 125–140. <https://doi.org/10.1027/1864-1105/a000118>
- [4] Jeremy Boy, Françoise 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* (Seoul, Republic of Korea) (CHI '15). Association for Computing Machinery, New York, NY, USA, 1449–1458. <https://doi.org/10.1145/2702123.2702452>
- [5] Judd D Bradbury and Rosanna E Guadagno. 2020. Documentary narrative visualization: Features and modes of documentary film in narrative visualization. *Information Visualization* 19, 4 (2020), 339–352.
- [6] Rick Busselle and Helena Bilandzic. 2009. Measuring narrative engagement. *Media psychology* 12, 4 (2009), 321–347.
- [7] Eun Kyoung Choe, Bongshin Lee, et al. 2015. Characterizing visualization insights from quantified selfers’ personal data presentations. *IEEE computer graphics and applications* 35, 4 (2015), 28–37.
- [8] Shauna Concannon, Natasha Rajan, Parthiv Shah, Davy Smith, Marian Ursu, and Jonathan Hook. 2020. *Brooke Leave Home: Designing a Personalized Film to Support Public Engagement with Open Data*. Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376462>
- [9] Irene Costera Meijer. 2013. When news hurts: The promise of participatory storytelling for urban problem neighbourhoods. *Journalism Studies* 14, 1 (2013), 13–28.
- [10] Auriol Degbelo, Carlos Granell, Sergio Trilles, Devanjan Bhattacharya, Sven Casteleyn, and Christian Kray. 2016. Opening up smart cities: citizen-centric challenges and opportunities from GIScience. *ISPRS International Journal of Geo-Information* 5, 2 (2016), 16.
- [11] Steven Drucker, Samuel Huron, Robert Kosara, Jonathan Schwabish, and Nicholas Diakopoulos. 2019. Communicating data to an audience.
- [12] Brent Dykes. 2019. *Effective data storytelling: how to drive change with data, narrative and visuals*. John Wiley & Sons, 111 River Street Hoboken, NJ 07030 United States.
- [13] Jose García-Avilés. 2020. *Reinventing Television News: Innovative Formats in a Social Media Environment*. Springer, Dordrecht, 143–155. [https://doi.org/10.1007/978-3-030-36315-4\\_11](https://doi.org/10.1007/978-3-030-36315-4_11)
- [14] Muki Haklay. 2013. *Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation*. Springer, Gewerbestrasse 11, 6330 Cham, Zwitterland, 105–122. [https://doi.org/10.1007/978-94-007-4587-2\\_7](https://doi.org/10.1007/978-94-007-4587-2_7)
- [15] 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* (SEOUL, Republic of Korea) (TVX '18). Association for Computing Machinery, New York, NY, USA, 43–55. <https://doi.org/10.1145/3210825.3210826>
- [16] David Li, Eric Lee, Elijah Schwelleng, Mason G. Quick, Patrick Meyers, Ruofei Du, and Amitabh Varshney. 2020. MeteoVis: Visualizing Meteorological Events in Virtual Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI EA '20). Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3334480.3382921>
- [17] Nan Li and Amanda L. Molder. 2021. Can scientists use simple infographics to convince? Effects of the “flatten the curve” charts on perceptions of and behavioral intentions toward social distancing measures during the COVID-19 pandemic. *Public Understanding of Science* 30, 7 (2021), 898–912. <https://doi.org/10.1177/09636625211038719> arXiv:<https://doi.org/10.1177/09636625211038719> PMID: 34448434.
- [18] Roberto Martínez-Maldonado, Vanessa Echeverría, Gloria Fernández Nieto, and Simon Buckingham Shum. 2020. *From Data to Insights: A Layered Storytelling Approach for Multimodal Learning Analytics*. Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3313831.3376148>
- [19] Eef Masson and Karin van Es. 2020. *5. Rain on your radar : Engaging with weather data visualizations as part of everyday routines*. Amsterdam University Press, Nieuwe Prinsengracht 89, Amsterdam, North Holland, The Netherlands, 77–94. <https://doi.org/10.1515/9789048543137-009>
- [20] Karen K Meltzer. 2017. Family activity model: Crossroads of activity environment and family interactions in family leisure. *Leisure Sciences* 39, 5 (2017), 457–473.
- [21] Mithra Moezzi, Kathryn B Janda, and Sea Rotmann. 2017. Using stories, narratives, and storytelling in energy and climate change research. *Energy Research & Social Science* 31 (2017), 1–10.
- [22] Dennis Nguyen. 2021. Mediatisation and datafication in the global COVID-19 pandemic: on the urgency of data literacy. *Media International Australia* 178, 1 (2021), 210–214.
- [23] Jonathan A Obar. 2020. Sunlight alone is not a disinfectant: Consent and the futility of opening Big Data black boxes (without assistance). *Big Data & Society* 7, 1 (2020), 2053951720935615.
- [24] William Odom, Ron Wakkary, Youn-kyung Lim, Audrey Desjardins, Bart Hengeveld, and Richard Banks. 2016. From Research Prototype to Research Product. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 2549–2561. <https://doi.org/10.1145/2858036.2858447>
- [25] Deokgun Park, Mohamed Suhail, Minsheng Zheng, Cody Dunne, Eric Ragan, and Niklas Elmqvist. 2022. StoryFacets: A design study on storytelling with visualizations for collaborative data analysis. *Information Visualization* 21, 1 (2022), 3–16. <https://doi.org/10.1177/14738716211032653> arXiv:<https://doi.org/10.1177/14738716211032653>
- [26] Charles Perin. 2018. Cycles and (A)Symmetry Exploring the Design of Shareable Personal Visualizations. In *IEEE VIS 2018 Arts Program Annotated Portfolios*. IEEE VIS 2018 Arts Program Annotated Portfolios, Berlin, Germany, 1–10. <https://hal.archives-ouvertes.fr/hal-01910357>
- [27] Annemarie Quispel, Alfons Maes, and Joost Schilperoord. 2016. Graph and chart aesthetics for experts and laymen in design: The role of familiarity and perceived ease of use. *Information Visualization* 15, 3 (2016), 238–252.
- [28] Byron Reeves, Leo Yeykelis, and James J Cummings. 2016. The use of media in media psychology. *Media Psychology* 19, 1 (2016), 49–71.
- [29] Nathalie Henry Riche, Christophe Hurter, Nicholas Diakopoulos, and Sheelagh Carpendale. 2018. *Data-driven storytelling*. CRC Press, 6000 Broken Sound Parkway Northwest, Suite 300 Boca Raton, FL 33487 United States.
- [30] Nina Schnyder, Stéphanie Baggio, Astrid Rossegger, Jérôme Endrass, Leonel Cunha Gonçalves, and Jennifer Inauen. 2021. Adherence to public health measures during the COVID-19 pandemic: the pivotal role of behavioural science.
- [31] Edward Segel and Jeffrey Heer. 2010. Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics* 16, 6 (2010), 1139–1148.

- [32] Ian Spence and Peter Krizel. 1994. Children’s perception of proportion in graphs. *Child Development* 65, 4 (1994), 1193–1213.
- [33] Anselm Strauss and Juliet M Corbin. 1997. *Grounded theory in practice*. Sage, Thousand Oaks, California.
- [34] Carlos Toural-Bran, Ángel Vizoso, Sara Pérez-Seijo, Marta Rodríguez-Castro, and Maria-Cruz Negreira-Rey. 2020. *Information Visualization in the Era of Innovative Journalism*. Routledge, Milton Park, Abingdon-on-Thames, Oxfordshire, England, UK.
- [35] Milka Trajkova, A’aeshah Alhakamy, Francesco Cafaro, Rashmi Mallappa, and Sreekanth R. Kankara. 2020. *Move Your Body: Engaging Museum Visitors with Human-Data Interaction*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376186>
- [36] Michelle Wilkerson, William Finzer, Tim Erickson, and Damaris Hernandez. 2021. Reflective Data Storytelling for Youth: The CODAP Story Builder. In *Interaction Design and Children* (Athens, Greece) (IDC ’21). Association for Computing Machinery, New York, NY, USA, 503–507. <https://doi.org/10.1145/3459990.3465177>
- [37] Gerard Wilkinson, Tom Bartindale, Tom Nappey, Michael Evans, Peter Wright, and Patrick Olivier. 2018. Media of Things: Supporting the Production of Metadata Rich Media Through IoT Sensing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI ’18). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3173780>
- [38] Annika Wolff, Daniel Gooch, Jose Montaner, Umar Rashid, and Gerd Kortuem. 2016. Creating an Understanding of Data Literacy for a Data-driven Society. *The Journal of Community Informatics* 12 (08 2016), 1–10. <https://doi.org/10.15353/joci.v12i3.3275>
- [39] Arch G Woodside. 2010. Brand-consumer storytelling theory and research: Introduction to a Psychology & Marketing special issue.
- [40] Niels Wouters, John Downs, Mitchell Harrop, Travis Cox, Eduardo Oliveira, Sarah Webber, Frank Vetere, and Andrew Vande Moere. 2016. Uncovering the Honeypot Effect: How Audiences Engage with Public Interactive Systems. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems* (Brisbane, QLD, Australia) (DIS ’16). Association for Computing Machinery, New York, NY, USA, 5–16. <https://doi.org/10.1145/2901790.2901796>
- [41] Jetmir Xhembulla, Irene Rubino, G. Malnati, and Claudia Barberis. 2014. Intrigue at the Museum: Facilitating Engagement and Learning through a Location-based Mobile Game. In *Proceedings of the 10th International Conference on Mobile Learning 2014, ML 2014*. *Proceedings of the 10th International Conference on Mobile Learning 2014, ML 2014* 10, 1–10.
- [42] Yi Zhu, Ingrid Heynderickx, and Judith A Redi. 2015. Understanding the role of social context and user factors in video quality of experience. *Computers in Human Behavior* 49 (2015), 412–426.