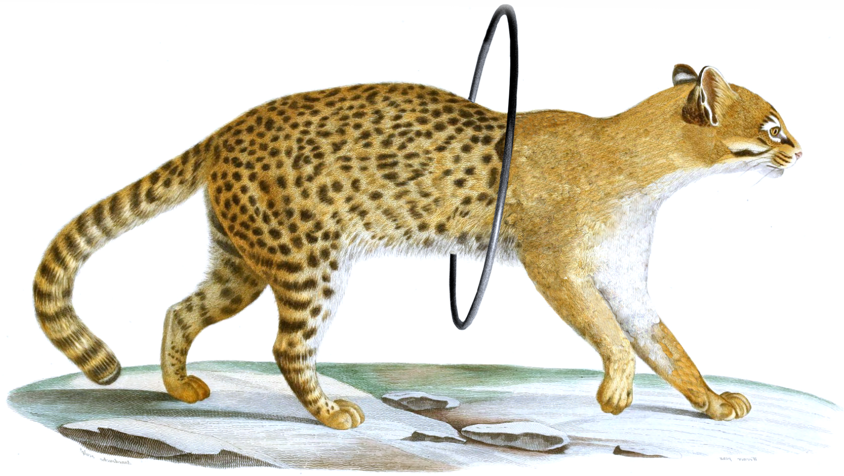
University of Applied Sciences and Arts Northwestern Switzerland
School of Engineering

Lost in transition

The effects of animation on the understanding of transitions in narrative visualization

Jonas Oesch

Master’s Thesis



Supervised by Prof. Dr. Simon Schubiger

MSc. in Engineering

University of Applied Sciences and Arts Northwestern Switzerland

August, 2019

Abstract

The research on data visualization has traditionally focused on maximizing the throughput of information through the reader's visual system. This is useful for researchers who need to explore large amounts of data but it is very strenuous. Narrative visualization, on the other hand, is a new approach to visualization that uses storytelling techniques to reduce the mental burden of the reader. One of these techniques is to guide the reader by organizing multiple charts in a sequential, narrative order. While this type of presentation splits information processing into manageable chunks, the reader now needs to understand how the individual charts are related. In practice, animated transitions are often used for this. But existing research indicates that animation might not be an effective way to present abstract relationships. This thesis aims to find out if animated transitions have a positive impact on the readers understanding of narrative visualization.

Based on a review of the literature on transition understanding and animation, we introduce the concept of *characters* to narrative visualization. We propose that readers understand transitions by comparing how characters and their setting changes between two charts. The strengths of animation were evaluated based on how they would serve this process. This has led to the formulation of several hypotheses.

An experiment presented 8 transitions, half of them animated, to 56 participants to test these hypotheses. The results indicate that animated transitions do not lead to higher overall understanding. They thus confirm existing research on the use of animation within the context of narrative visualization.

Acknowledgements

Many thanks go to Rebecca for her patience with my absent-mindedness and for always being cheerful.

I thank Simon Schubiger for all the philosophical discussions which I highly enjoyed.

I thank Nicole Jardine for the many exchanges and her input on the experiment. I equally thank Steven Franconeri, Cindy Xiong, and Griffin Bianchi for their inputs.

I thank Thomas Nemecek, Katharina Besteler, Madlaina Kalunder and Rebecca Nemecek for their feedback on the writing.

I thank the fantastic 5.2C19 office, Andrea, Cloe, Filip, Simon, and Simon for all the good times.

I thank the many people who have contributed to this work by looking at the designs and sharing their ideas who will sadly remain unnamed.

Finally, I thank God, our Father, for His wonderful and exciting creation that gave me so much to explore and for sustaining me through all highs and lows.

Contents

1. Introduction

For many people, when they hear the word “data”, it evokes feelings of something rather dry, even lifeless. They might have heard that the data generated each year is continuing to grow exponentially each year [1]. They might have heard of the many marvels that data will deliver to humanity. But they simply can not relate. Data is something for the specialists who have a special trait of character, something that escapes the regular person.

This view of things is reinforced by most of the research that is being conducted around data. Much of it is focused on how to store, process and interpret the ever-larger amounts of data. And while these researchers push the boundaries of what is technologically possible, human capabilities remain painfully limited within narrow boundaries. It is obvious how this situation leads some to conclude that we need to either augment the human brain through technology [2] or to remove the human from the loop entirely [3].

User interface design in general and data visualization, in particular, have traditionally taken the role of the mediator between the two worlds. Visualization research has extensively explored how to maximize the data that can be communicated to humans. They have faced limitations like available pixels on the screen and the perceptual and cognitive abilities of people. Much of the thinking in the field is focused on not wasting these precious resources. It is epitomized in recommendations like “maximizing the data-to-ink ratio” [4]. and in observations like this: “The visual system provides a very high-bandwidth channel to our brains” [5, p. 6].

We argue that it is precisely this narrow focus on maximizing data throughput that has alienated the regular person from data. This is because feeding a person the maximum amount of data possible comes at a cost. While the visual system and the brain might be able to process a surprisingly high amount of information, it puts a lot of strain on them. A researcher who is driven by the prospect of finding answers to his questions in the data will probably have a high tolerance to put up with this. But even for them, data analysis is a fatiguing task.

Visualization that is intended to communicate therefore can ill afford to burden the recipient the task of exploration. And while this might seem obvious, the distinction only emerges slowly in the visualization community. A recent tweet from John Stasko, a visualization researcher for over 30 years, shows this [6]:

> I grow increasingly convinced every day that data visualization for analytical, exploratory purposes and data visualization for communicative, presentational purposes are more different than most people think.

1.1. Narrative visualization

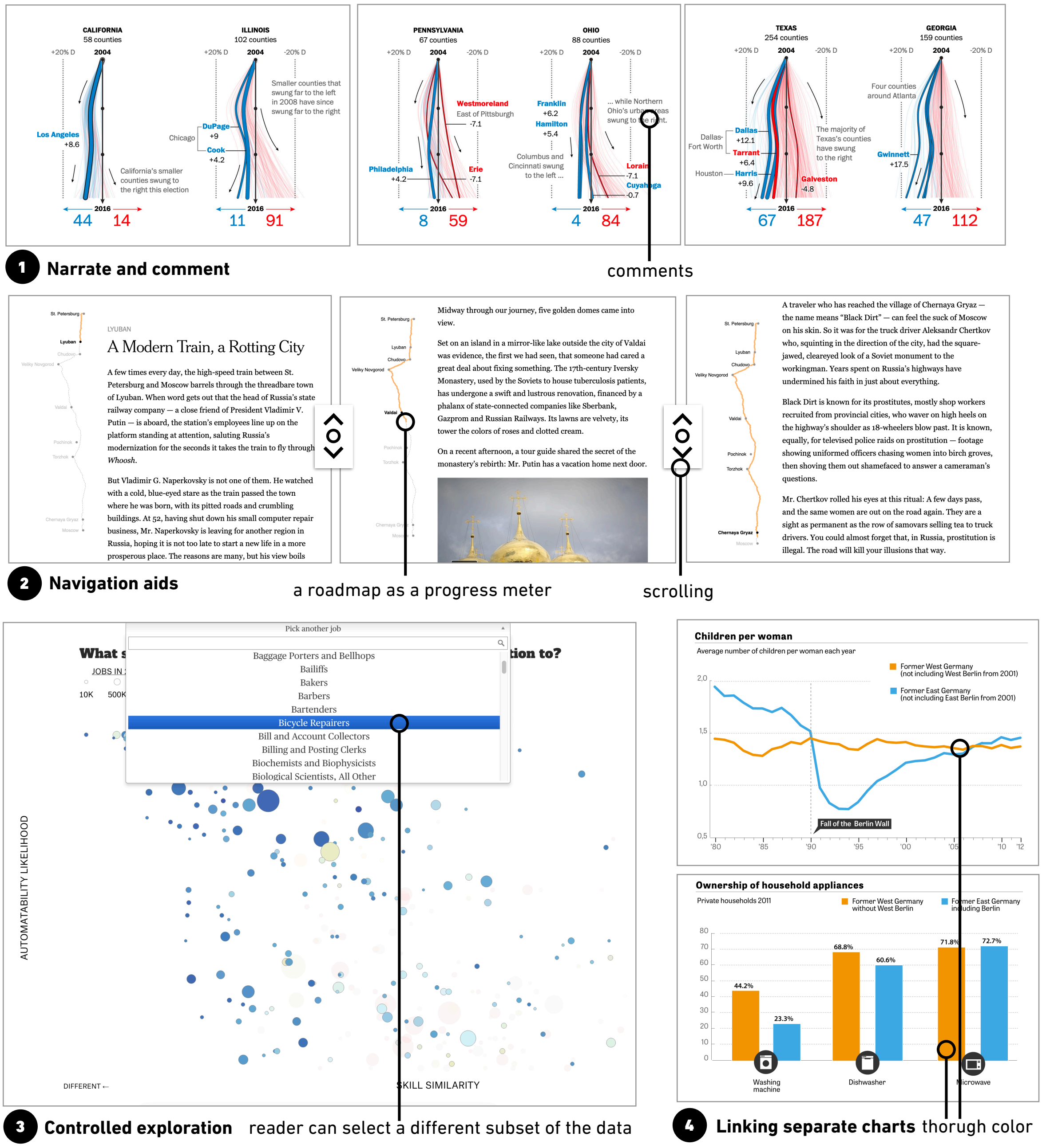


Figure 1: 1. The Washington Post uses annotation extensively to the explain the changing voting behaviors of counties [7]. 2. In this story about rural Russia [8], a map is used as a navigational aid. 3. At the end of an article on the job propsects of truck drivers that will be replaced by self-driving cars[9], the reader can explore other jobs. 4. From an article on still existing differences between Eastern and Western Germany [10]. The two regions are consistently identified by their color.

After a few years of experimentation with including interactive, explorative visualization in their stories, multiple online newspapers have found that users do not click on them, except when the authors make it abundantly clear why they should interact [11], [12]. Archie Tse from the New York Times puts it like this [12]:

> If you make the reader click or do anything other than scroll, something spectacular has to happen.

A study by Boy et al. [13] titled "Does it Engage Users to Explore Data?" came to a similar conclusion: Participants preferred to gain information from a textual narrative rather than from an interactive visualization. Their preference changed only when the authors showed an animated mouse cursor that hovered over the visualization and showed how participants would quickly get the answer through interaction. These findings indicate that readers prefer a high amount of guidance in presentations of data because it removes the burden of exploration from them.

A recent field of research, called *narrative visualization*, is precisely concerned with the question of how to best guide readers when presenting data. Narrative visualizations on the surface share many traits with conventional visualization but they but they have four distinct characteristics that are all related to guiding the reader [14, p. 85] (see figure 1): they **narrate and comment** a visualization through text, audio, and annotations; they provide **navigation aids** like timelines, maps, breadcrumbs, etc.; they provide **controlled exploration** through embedded interactive visualization; and they **link separate charts** through color, animation and/or interaction

As the name implies, narrative visualization puts charts into a narrative sequence. Such logical sequences are highly desirable when communicating as they lead to better understanding and higher recall of the information presented [15]. But simply putting charts into a narrative. logical sequence is not enough. The author needs to somehow communicate how the individual charts are related, to link them [14, p. 92].

This can be done via an accompanying narrative text, but visual methods like color and animation have also been used in practice (see number 4 in figure 1). But while animated transitions between charts have quickly gained popularity in recent times (sec. 4.3), existing research does not necessarily indicate that animated transitions provide any benefits [16], [17]. This motivates the present research on animated transitions in narrative visualization.

1.2. Problem definition

The main objective of the present thesis is to find out if animated transitions help readers understand the relationship between two charts in the context of narrative visualization.

To answer this question, the following sub-questions need to be answered first:

- How do readers interpret transitions between charts in narrative visualizations?

- How could animation support this process?

To create an experiment to answer the main research question, a software needs to be developed that supports the creation of narrative visualizations with animated transitions.

1.3. Thesis outline

The thesis is organized in five parts:

1. The first part presents the theoretical background on transitions and animation. It shows the strengths and weaknesses of existing classifications of transitions and discusses how readers interpret transitions. The review of the literature on animation presents five strengths of animation and discusses two important principles for animation design.
2. The second part introduces a reconceptualization of the existing literature to operationalize it for animated transitions in narrative visualizations. It goes on to show that the reconceptualization is applicable to narrative visualization from practice and reformulates the existing classifications. Based on this reformulation, it discusses how the strengths of animation presented in part one might be applied to different transition types. This leads to the formulation of six research hypotheses.
3. The third part presents an experiment that was conducted to answer the research hypotheses in part two.
4. The fourth part presents a declarative syntax based on the reconceptualization in part two that was used to design the narrative visualizations for the experiment in part three.
5. The fifth part presents our conclusions and possibilities for future work.

Part I.
Theoretical background

2. Transitions

Presenting charts in a logical, narrative sequence seems to be preferred by readers [18]. This implies that there is some information in the sequencing that is not present in the charts themselves. But what kind of information? This part reviews how previous work classifies transitions according to the kind of information that they transport. It then takes a deeper look into how readers interpret transitions to gain a better conceptual understanding of transitions and what makes them difficult. It then applies this conceptual understanding to examples of narrative visualization from practice to reformulate and synthesize the existing classifications. Finally, it discusses the implications of this sharpened understanding for designing animated transitions.

2.1. Existing classifications of transitions

Multiple authors have compiled classifications of transitions between charts. In this section, three different approaches to classifying transitions will be presented: visualization-focused, story-focused and inductive.

2.1.1. Visualization-focused classifications

The most common approach is to think about transitions is in terms of visualization parameters [19]–[22]. Any visualization can be viewed as the result of a series of parametrizable transformations applied to the underlying data set [23]. The transition can, therefore, be described in terms of parameter changes as long as the underlying data does not change.

Through this approach, four authors identify a total of 9 transition types. Because their approach is very similar, their classes have a high overlap and will be presented together. Figure 1 maps the names the authors have used to the "generic" names that we have used in their description:

Table 1: Comparison of transition types commonly identified by different authors for different purposes. The table shows that the agreement is relatively high for transitions that can be defined in terms of visualization parameters. What is often missing is the case where the underlying data changes.

	Heer et al. [19]	Yi et al. [20]	Fisher et al. [21]	The Space Between [22]
Pan	View transformation	Explore	Change the view	Pan & Zoom
Zoom	View transformation	Abstract / Elaborate	Change the view	Pan & Zoom
Filter	Filtering	Filter	Filter the data	-
Reorder	Ordering	Reconfigure	Reorder the data	Sort order
Substrate transformation	Substrate transformation	Reconfigure	Change the charting surface	-
Visualization change	Visualization change	Encode	Change the representation	Graph type
Data schema change	Data schema change	-	Change the data	Variable transition
Data change	Timestep	-	Change the data	Data transition

- **Pan**: Moving a fixed image below a smaller viewport to see one part at a time. The most famous example of this is Google Maps [24].
- **Zoom**: Geometric zooming in and out of a fixed image. This transition too is used in Google Maps [24].
- **Filter**: Hiding or showing elements based on some criteria. When reading a visualization on cars, filtering might, for example, hide all the data on American cars.
- **Reorder**: Reordering axes according to a different criterion. Countries that are displayed in alphabetical order might, for example, be reordered according to their GDP.
- **Substrate transformation**: Any distortion to the scale that is not addition or multiplication of a constant. For example a log transform or a lens-effect that resizes only a part of the coordinate system. The latter being something that is not often seen in real-world applications but seems to

be a darling of visualization researchers [5, pp. 327–338].

- **Visualization change**: Showing the same data with a different visualization idiom. For example, showing a distribution as a boxplot to indicate mean, quantiles and outliers first and then transforming it into a histogram to give a better sense of its shape.
- **Timestep**: Showing the same visualization for different points in time. For example, the population pyramid of Switzerland for 1960 and 2010 to show how it has shifted to older people.
- **Data schema change**: Mapping a different data dimension to one of the visual variables (position, color, etc.). A visualization might, for example, show the use of certain plant variants over time and show the occurrence of pests over the same period in the next state.
- **Data change**: When new data is shown but the visualization and the axes stay the same. This is the generalization of the *timestep* transition.

The strength of this kind of classification is, that types can relatively easily and coherently be determined as the high overlap between different authors shows. Its weakness is that it characterizes a transition by its technical properties and not based on how the reader will interpret it, which makes it less useful for designers of narrative visualization. Other fields, especially the research on comics have done much more work in this regard.

2.1.2. Story-focused classification

The most-cited classification of transition types in comics is the one by McCloud [25]. It is also commonly referred to in narrative visualization [18], [26]–[29] because comics are similar to narrative visualization in that they are *visual* and *sequential* [25, p. 16]. McCloud describes six types of transitions between comic panels (see figure 2):

1. **Moment-to-moment**: Stepwise movement through time.
2. **Action-to-action**: Movement through time-based on meaningful events.
3. **Subject-to-subject**: Switch from one subject to another
4. **Scene-to-scene**: Change in location or a long timespan passing.
5. **Aspect-to-aspect**: Exploring a scene.
6. **Non-sequitur**: Changes without an obvious relationship.

In McCloud's classification, a transition is characterized by a change in either *subject*, *scene* and/or *time*.

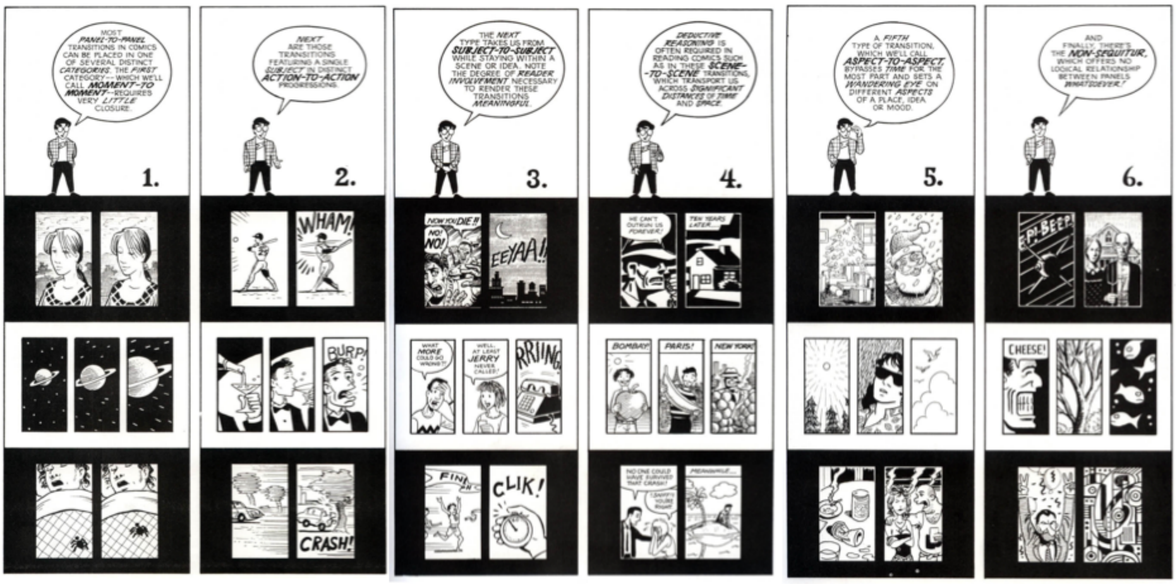


Figure 2: The six transition types for comics proposed by McCloud based on the concepts of *subject*, *scene* and *time* [25].

Cohn [30] refines on McCloud's transition types by introducing hierarchy to the interpretation of subject, scene and time. He argues that readers will first try to find the same subject in the subsequent panel. They will then move their attention to changes in the scene. Finally, they will shift their focus to changes in time. Cohn uses examples to show how a different order leads to interpretations that run counter to typical reading habits [30, p. 143].

Subject, *scene* and *time* are reminiscent of journalisms famous *who/what*, *where* and *when* [31]. Such proximity to journalism indicates that McCloud's transition types might be useful outside of the domain of comics. While McCloud's transitions based on *subject*, *scene* and *time* might be useful in traditional narratives that talk about people and events, it is not directly obvious how to apply them to abstract visualizations [28, pp. 68–82].

2.1.3. Inductive classification

A bottom-up approach to transition types can be found in Hullman et al. [32]. The authors reviewed 42 narrative visualizations and identified 12 different transition types which they grouped into six categories as presented in figure 3:

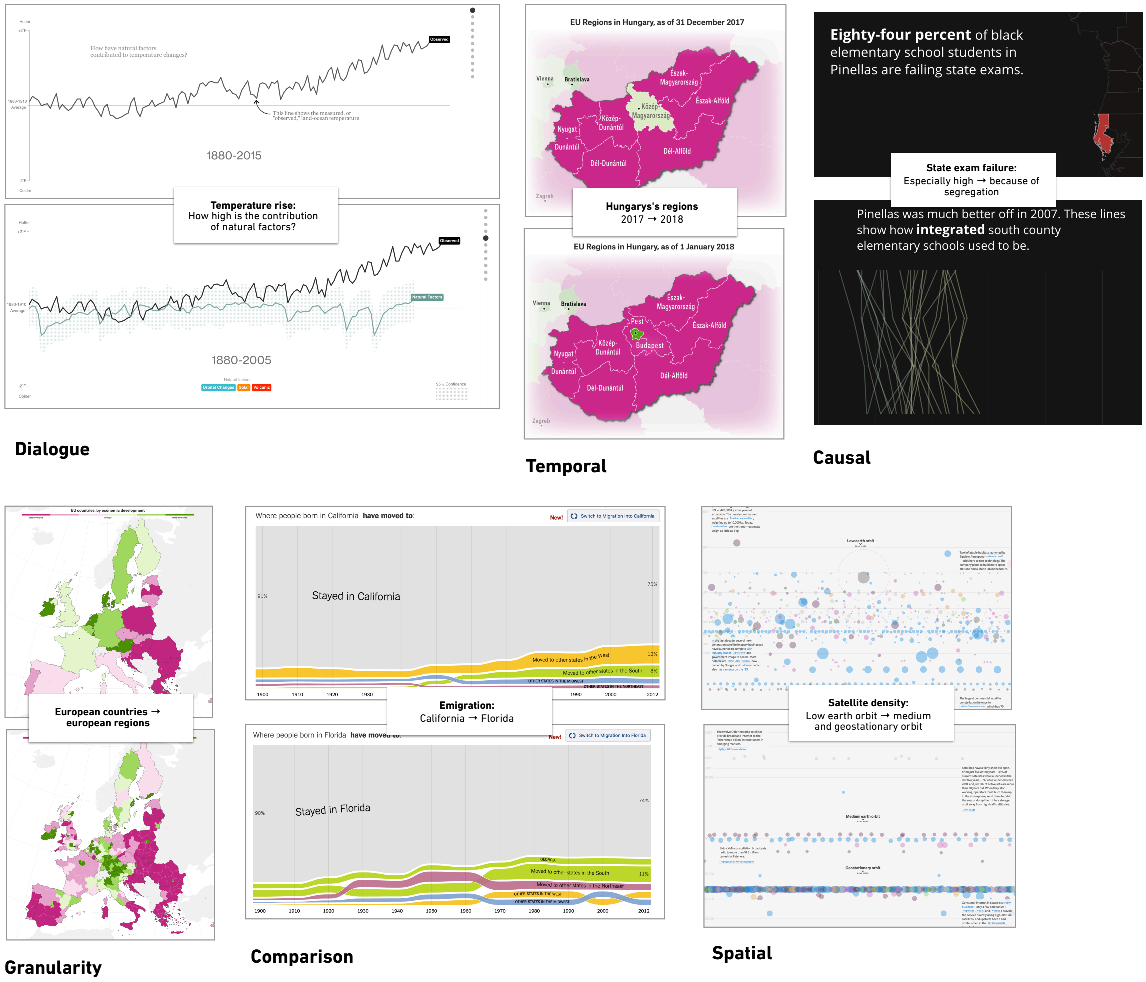


Figure 3: The six categories of transitions identified by Hullman et al. based on an analysis of 42 narative visualizations [32]. The sources of the visualizations used to illustrate the transitions are (in order): [33], [34], [34]–[37]

Dialogue: A question is asked or implied in one chart which is being answered in the next. One chart shows a time series of world temperature and asks how natural factors have contributed. The next chart shows how natural factors like volcanos have influenced global temperatures over time (see figure 3 on the top).

Temporal: The two charts show the same data for different points in time. For example how the regional subdivision of Hungary has changed between 2017 and 2018 (see figure 3 on the top).

Causal: One chart shows an effect and the next chart presents the cause of the effect. A chart showing a school district with high rates of failures and the next showing how high failure rates have been caused by higher segregation, for example (see figure 3 on the top).

Granularity: When one chart shows a higher amount of detail than the other. One example is a chart that shows EU countries on a map and another chart that shows how they are split into regions (see figure 3 on the bottom).

Comparison: Either the dependent or an independent variable changes between the charts. The transition is called a measure walk when the measure is being changed. For example, showing a map of the impact of climate change on flooding risk and a map of the impact on crop shortage (where the respective impacts are the dependent variables). It is called a dimension walk when one of the independent variables changes. For example, a series of sorted streamgraphs that show emigration patterns for different states (where the state is an independent variable) (see figure 3 on the bottom).

Spatial: When the final state is spatially close to the initial state, as visible in figure 3 on the bottom, where the charts show different heights of the earth's orbit.

Hullman et al. further distinguish between *implicit transitions* that can only be interpreted by a reader and *explicit transitions* that could potentially be inferred from changes in the attributes visualized (compare to section 2.1.1). Interestingly they have found very little occurrences of the two explicit transitions in their analysis (see figure 4).

Category	Category prevalence	Transition Types	Total
Dialogue	17%	Question & Answer	4
		Who, What, When, Where, Why, How	3
Temporal	88%	Simple chronological	29
		Reverse chronological	11
		Future chronological	12
Causal	24%	Explicit Cause	7
		Alternative Reality	3
Granularity	71%	General to specific	28
		Specific to General	16
Comparison	64%	Dimension Walk	20
		Measure walk	19
Spatial	24%	Spatial Proximity	10

Figure 4: Transition types identified by Hullman et al. [32]. The explicit *dialogue* and *causal* transitions are quite rare. Also *spatial* transitions did not occur very often.

The strength of Hullman et al.'s classification is, that it is directly linked to real-world examples of narrative visualization. But because it has no underlying conceptual understanding, it is often not clear how to classify a transition according to their taxonomy.

2.1.4. Limitations of existing classifications

The *visualization-oriented* (sec. 2.1.1) and the story-oriented (sec. 2.1.2) classifications both provide a conceptual understanding of transitions based on the question: "What changes?". In the first case, it is one of the visualization parameters. In the second case, it is either *subject*, *scene* or *time*. This allows for the construction of theory around these transitions [19], [30] because there is clarity about the basic concepts. But these classifications can not be directly applied to narrative visualization because they were intended for other contexts.

Hullman et al.'s inductive approach [32] is interesting because it treats real-world examples of narrative visualization. But it has no theoretical concepts linked to it. It can, therefore, not be used to understand how readers understand transitions and how animation might support it.

The following section will, therefore, discuss the existing literature on how readers perceive transitions. This review will form the basis of a reformulation of Hullman et al.'s transition types in terms that link it to a theory. This theory, in turn, will help to discuss how animation can be used to guide readers.

2.2. How readers interpret transitions

In the previously discussed classifications, transitions are often classified according to the question: "What changes?". This implies that readers make sense of transitions through comparison. This section, therefore, gives a brief overview of the relevant work on comparison in visualization. Based on this, it discusses what makes transitions difficult for readers. It then introduces findings from cognitive load theory that help in understanding how these difficulties are linked to the cognitive abilities of readers.

2.2.1. Comparison to understand transitions

Readers find and interpret relationships within and between charts by comparing targets from the charts in their head. This implies that interpreting relationships is related to the number of targets that can be held in the reader's mind.

According to Brehmer et al. [38], comparison is a fundamental task in visualization. Comparison is distinguished from other tasks like *identify* and *summarize* (see figure 5) by the number of *targets* it operates on. An *identify*-task uses a single target while a *summarize*-task uses all available targets. But the *compare*-task requires multiple targets to compare them.

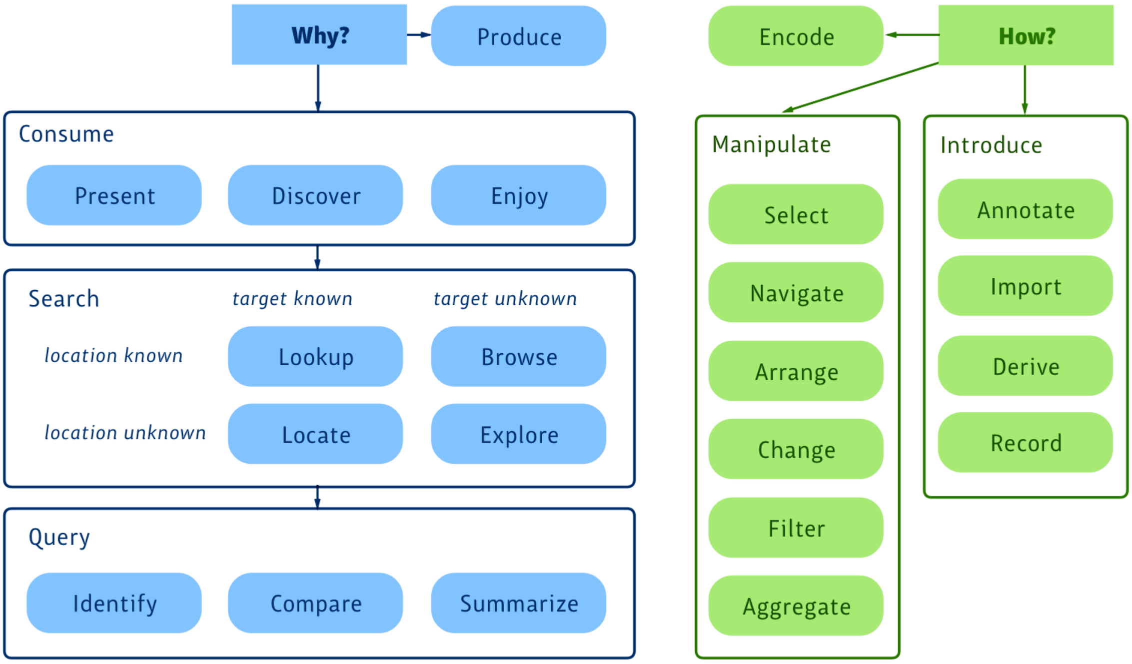


Figure 5: A topology of visualization tasks which synthesizes a large body of previous literature has been proposed by Brehmer et al. [38]

Before he can to compare targets, the reader needs to *search* for an initial target (see figure 5). In narrative visualization, this will usually be done by *locate* to find targets that are mentioned in the text. But this is not the only possibility:

Explicit identification Is aided by the author through labels, annotations, colors, and narrative. An example can be seen in the first point of figure 1.

Identification through visual statistics This uses the ability of the visual system to quickly and quite accurately calculate certain statistical properties of an image (see figure 5.1.3). A viewer can, for example, see in an instant if a certain region contains more orange than blue dots (mean). He can also quickly identify groups of points that are similar (clustering). Viewers will be able to tell if circles get bigger from left to right (trends). Or if a point has a different color from all the outers (outlier). All of these values can then serve as a target for comparison.

Implicit identification In many cases, external knowledge or personal interest of the viewer will lead to the identification of a target [39]. Often, readers will compare what they see to what they would have expected to see. Or a person from Brooklyn, for example, will tend to be more interested in the success and failure history of the Brooklyn Nets than that of the Warriors from California (see the third example in figure 1). This is one of the reason for authors of narrative visualization to include *interactive visualization*.

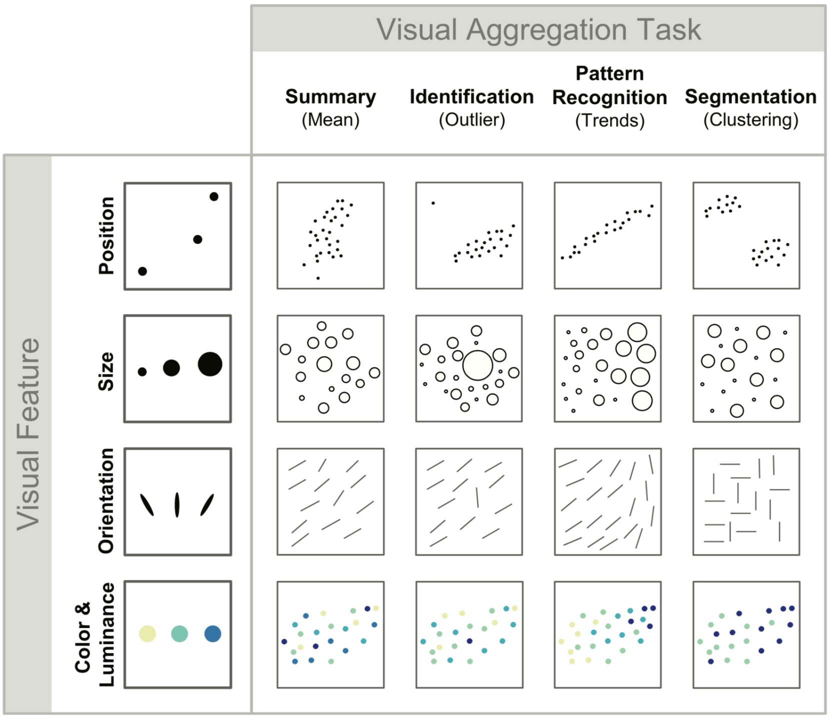


Figure 6: The visual aggregation tasks from [40] show how comparison targets can by identified through visual statistics.

When a target or a set of targets have been identified in one of the charts, the reader proceeds to *locate* a corresponding target in the other chart. When all the targets are available, they are compared in the reader's mind. This often requires a considerable mental effort, which we will call the *transition cost*.

2.3. Transition cost

The transition cost describes how much mental effort it takes to do the necessary comparison(s) to understand a transition. Gleicher [39] describes three factors that make comparisons challenging:

- A large number of targets.
- Large targets. For example when comparing two very long time series.
- Complex relationships. For example when there is not a one-to-one relationship.

Cognitive load theory predicts that a maximum number of 9 targets can be held in working memory for processing [41], [42]. Depending on the size of the targets and the complexity of the relationships, this number can get as low as 1 [43]. When the number of targets that are needed for a comparison exceeds these limitations, the reader will forget some of them and the comparison fails [41].

When designing narrative visualization, the cost of the transitions should ideally stay within these boundaries to make the reading fluent. Animation is often proposed as a means to make transitions easier, but it is often not clear how exactly this works. The next chapter will, therefore, present the relevant research on the use of animation in visualization.

3. Animation

3.1. Strengths of animation

Animation has been portrayed to provide many benefits to UX-design in general and visualization design in particular. Chevalier et al. [44] have identified 23 different "purposes" of animation in visualization. Things like "Staying oriented during navigation" or "Hooking the user" which are finally classified into five meta-classes. Although this overview shows the popularity of animation in today's practice, it does not discuss any evidence that animation *actually* serves the given purposes.

To understand why designers might think that animation helps in these specific circumstances, we have mapped some lower-level strengths of animation described in Heer et al. [19] to all each of the purposes (see figure 7).

Keeping context	Staying oriented during navigation	A	Maintaining up to date	A
	Supporting tracking during layout changes	A	Replaying history, summarizing	A
Teaching aid	Teaching a new representation of information	A		A
	Training, demonstration by example	A C	Illustrating an algorithm	A C E
	Explaining how a system works	A C	Affordance and preview	B C
User experience	Revealing/Hiding Content	A B	Keeping the user engaged	B D
	Providing a virtual tour	C	Providing visual comfort and aesthetics	D
	Making activity and progress visible	C B	Providing feedback of input mechanism	C D
	Hooking the user	B D	Conveying uncertainty or randomness	E
Data encoding	Revealing data relationships	A C	Encoding emotion	D
	Encoding object attribute	E		
Visual discourse	Supporting a narrative	A C	Persuading and convincing	C D
	Highlight content	B		

- A Object constancy
- B Attract attention
- C Communicate causality and intentionality
- D Engage
- E Data encoding

Figure 7: The purposes of animation found by Chevalier et al. [45] each employ one or more of the strenghts of animation that are described by Heer et al. [19].

3.1.1. Object constancy A

Object constancy describes the ability to identify two objects as being the same entity between two states. There is some evidence that when the reader has rapid visual access to the targets (less than 300ms [46]), changes will be detected before working memory is reached. The idea is that when one target is morphed into another, they are identified as being "the same" without having to do a comparison in working memory. This is probably the strength most often invoked for animation [19], [47]–[49].

3.1.2. Highlighting B

Movement very strongly attracts attention. This view is undisputed in the literature [5, p. 238] and there is some solid evidence that motion is more effective in highlighting elements than other visual means [50]. This ability to highlight elements can be very useful for storytelling. But it also poses the risk to distract the reader from important, non-moving elements of a visualization like axes or labels.

3.1.3. Communicate relationships C

In nature, we observe a certain smooth flow of cause and effect. When a rolling billiard ball hits another, the second ball will smoothly continue its transition. This leads to the intuition that smooth animation would imply some sort of cause and effect relationship (see section 3.2.1). This is true when depicting naturalist phenomena like the formation of lightning [51]. Other authors have found that animation also conveys causality in more abstract displays but that it does not perform better than symbolic depictions of causality [52].

Animation has also been reported to have modest benefits when depicting aggregation operations like summing [53].

3.1.4. Engage D

Multiple authors have remarked how readers were "excited" about animation in interfaces and that it thus led to higher engagement [16], [19]. The ability of animation to engage and excite is often presupposed based on anecdotal evidence. But few studies have used validated questionnaires to explicitly measure engagement. One study who did, found very mixed results [54]. In their findings animation only led to higher engagement when coupled with pictorial representations of data. The main factor that led to higher engagement was "Aesthetics". We suspect that in many cases there might be a "novelty" effect of animation that excites readers. There is little evidence that yesterdays fashionable animation still leads to higher engagement today. Furthermore, the novelty effect might lead readers and designers astray and they might confuse their positive feelings with actual benefits for understanding [26], [55].

3.1.5. Data encoding E

Similar to color or shape, motion can be used to encode data. Although the precision in interpreting it is very limited [5, p. 95]. This property of animation is not listed in Heer et al. [19] but we have decided to include it because it is well studied and forms the basis of multiple of the purposes listed by Chevalier et al [44].

While the listed strengths seem straightforward, the literature indicates that they are present only under certain circumstances.

3.2. Considerations when using animation

Animation does not automatically make transitions easier. Its effectiveness is highly related to two sets of principles: *congruence* and *apprehension*. When these principles are being violated, the effect of animation on the transition cost might be negligible or even negative.

3.2.1. Congruence

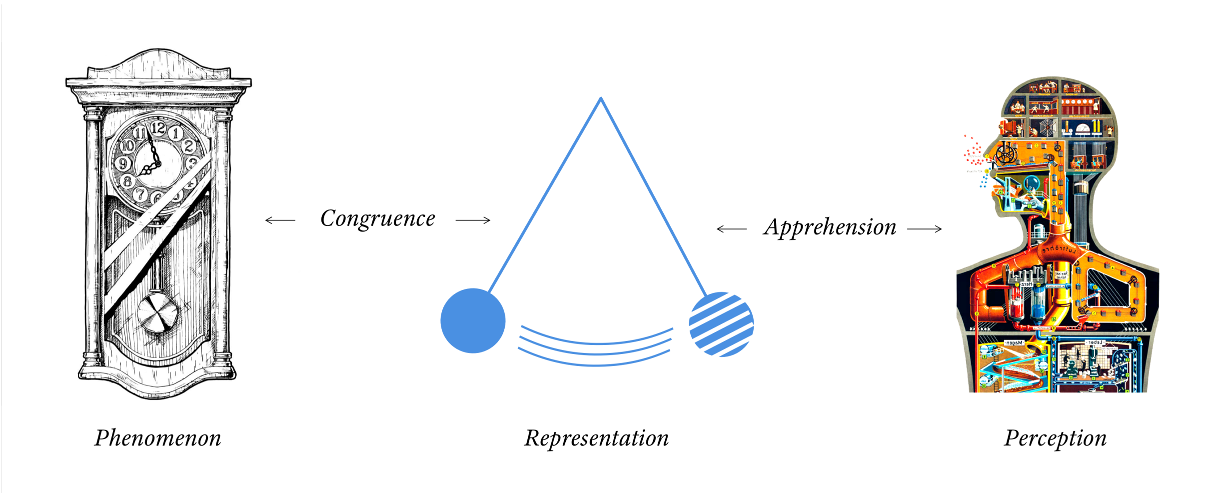


Figure 8: Congruence as defined by Tversky et al. [16] means that the representation should correspond to the concepts that are conveyed. Apprehension means that the representation should be accurately perceived.

The question if animation is a natural fit for certain types of transitions has been described as the principle of *congruence* [16, p. 247]:

> Effective graphics conform to the Congruence Principle according to which the content and format of the graphic should correspond to the content and format of the concepts to be conveyed.

According to this view, showing *change over time* would be a congruent use of animation. The reasoning is that people perceive change over time as smooth transitions in their everyday lives. One study accordingly found that animation reduced the perceived cognitive load when showing change in networks [56]. Other authors have not found such benefits [16] naming a lack of *apprehension* as the potential reason.

The second natural fit for animation are *viewport changes* like *pan* and *zoom*. These correspond to a person's natural movements through space, getting closer to objects, looking around, etc. Animated transitions have indeed been found to be superior to static transitions for *panning* [57] as well as *zooming* [58].

As previously discussed (3.1.3), animation has also been shown to be congruent for communicating causality and intentionality.

Interestingly, multiple authors found that animation seems to be far less effective when used with abstract representations than when it is used with iconic representations. Amini et al. [54] found that animated transitions in data visualization only led to higher engagement when coupled with iconic representations (see figure 9). A recent meta-analysis [17] of 50 studies on animation and its effect on learning found, that animation only benefitted knowledge acquisition when used with "iconic representations" of the phenomena the needed to be studied. No effect was found for abstract representations.

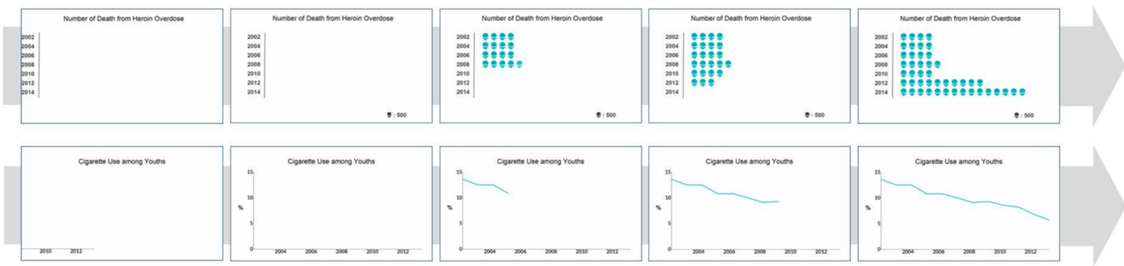


Figure 9: Amini et al. found that animation only led to higher engagement together with iconic representation of the data (on the top) while it did not differ from non-animated transitions when the representation was abstract (on the bottom) [54].

These findings indicate that the effectiveness of animation is strongly related to the amount of "realism" of the animation and even the visualization that it is being used on. But even when the use of animated transitions seems congruent, it may still fail. Tversky et al. cite a lack of *apprehension* for this [16].

3.2.2. Apprehension

Tversky et al. describe the apprehension principle like this [16, p. 247]:

> … graphics should be accurately perceived and appropriately conceived.

It is therefore related to the basic perceptual abilities of the reader. Heer et al.[19] define six design principles for animated transitions to improve apprehension: *similar transforms should be grouped*; *trajectories should be predictable*; *transformations* should be *simple*; to simplify complex transformations, *staging* should be used and the transition duration should be *long enough* to follow (but not longer). Although these principles seem to make sense intuitively, research on visual perception has shown to produce unintuitive results.

Visual tracking Research on visual tracking explores how people track objects when the objects are moving. Participants typically have to track the position of several individual objects (circles or rectangles) or a group of objects under different conditions. These experiments have revealed several, sometimes surprising, properties of the human perception:

Animation speed: Higher animation speeds do not strongly influence tracking ability [59].

Distance traveled: Longer distances traveled make objects harder to track [59].

Number of objects: In typical situations, four objects can be tracked in parallel [60].

Unpredictable paths: Have only a minimal impact on tracking performance [45].

Occlusion: Does not impair tracking when it can be interpreted as "one object disappears"behind another" [61].

Crowding: When objects that need to be tracked get close to each other, tracking performance is impaired [45].

Multiple objects are tracked as one convex area: Deformation of this area, as well as distractors entering the convex area, lead to impaired tracking performance [45], [60].

The above results show that object tracking happens through a complex interplay of the different levels of the visual system. While they give some indications on how the visual system processes information, the findings are hard to operationalize in animation designs because they will often conflict with the content that should be depicted. The *distance traveled,* for example, is often given by the data that is being depicted in the two charts and can not be influenced. The next section will discuss a few concrete propositions that have been made to increase the apprehension of animated transitions.

3.2.3. Animation techniques

A number of techniques have been proposed in the literature to increase the apprehension of animation. A brief discussion will show that in practice congruence and the limits of working memory seem to be more important than apprehension.

Staggered animation Because only a limited number of objects can be tracked simultaneously (see section 3.2.2), it has been proposed to animate objects in multiple steps (see figure 10).

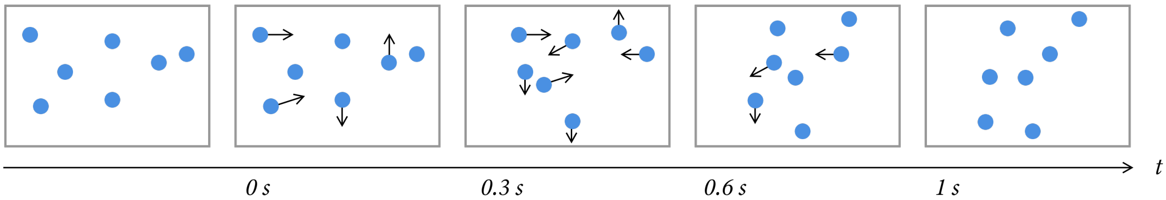


Figure 10: In staggered animation, some objects start to animate only with a delay.

Staggered animation delays the start time of animations incrementally for some objects. An example: out of 7 objects, 4 will start to move right away, 3 other objects will start to move only a bit later. A more complex example is shown in figure 10. The delay between the starts is called *dwell*. A dwell of zero means no staggering.

Staggered animation has been proposed to *reduce occlusion* as well as *complexity* during animation and therefore lead to a higher tracking performance [19] but interestingly the effect is negligible [45].

Staged animation Staged animation splits complex animation into multiple steps. An animation then who interpolates between positions as well as between two colors, might, for example, be split into two animation stages. One where the objects move from position a to position b. Another, where they change color.

Heer et al. explicitly recommend the use of staged animation [19] even though they found only modest benefits. In situations where staging led to complex and long animations, the authors found it even detrimental for tracking performance. This is also what cognitive load theory would predict because long, uninterrupted multi-stage animation will introduce elements into working memory but not leave the viewer time to organize them into long term memory. His working memory will therefore overflow and he will forget what he had seen before. This is called the *transient memory effect* [41].

Animation duration Multiple authors have observed that making animation slower is perceived negatively by readers [26], [57]. Others counter that very fast animation might lead to higher error rates

[19]. This assumption has been refuted by Franconeri et al. [59] who showed that higher speeds have no impact on tracking performance when the distance between objects stays large enough.

Direct manipulation Readers typically perform better at recall and problem solving, when they can control animation themselves. Multiple authors have shown this for cases where readers could start and stop animated transitions [62], [63]. But here too, it is hard to separate different effects. Tversky et al. point out, that interaction has been shown to benefit learning in itself and just happens to be coupled with animation often [16].

Easing John Lasseter in his classic paper on the principles of computer animation [64] describes that animators developed a preference for "slow-in-slow-out" over linear easing over time. This recommendation has been confirmed in more formal experiments [65]. Because "slow-in-slow-out" motions are closer to what we experience in nature, this type of easing can be seen as more *congruent* than other forms (see section 3.2.1).

3.2.4. The role of apprehension and congruence

The discussion of the research in the previous section demonstrates that clever technique to improve apprehension have most probably limited importance in practice. The proposed techniques often do not lead to the purported benefits, mostly because the easily lead to violations in other areas of apprehension, congruence or working memory.

The only two techniques that have consistently shown to be beneficial are direct manipulation and "slow-in-slow-out"-easing.

The principle of congruence, on the other hand, has been shown to have a profound impact on the effectiveness of animation. The closer the objects and the motion are to the real world, the more positive the effects of animation. This casts severe doubts on the usefulness of animation to support transitions between highly abstract visualizations.

3.3. Unfair comparisons

Further doubts on the reported benefits of animation come from the problem of *informational equivalence*. Some studies who have reported benefits of animated transitions have been criticized for involving "unfair comparisons" [16]. The animations presented contained information that was not easily deducible from the non-animated transitions. Tversky et al. [16, p. 251] put it like this:

> Showing that students learn material better when it is presented than when it is not presented should not be a goal of empirical research.

Fair comparisons, in contrast, need to identify *informationally distinct* states very clearly and show them not just in the animated scenario but also in all the scenarios they are compared to. Tversky et al. find that in these experiments, purported benefits of animation often fail to materialize.

Part II.
Reconceptualization of the theory

4. Perception-oriented classification

The discussion of the literature in the previous sections has shown that readers understand relationships by comparing targets in their working memory and that animation is more beneficial when it is applied to objects close to reality. It has also shown that none of the existing classifications could be used to link transition understanding to animation. This chapter, therefore, reconceptualizes the transition types seen in previous chapters so that it is explicit what the reader needs to compare. It also expresses the elements of comparison in terms of characters, attributes, and context to favor an interpretation that is closer to real-life concepts and should help to define congruent animations. This will lead to recommendations on how to use animation for the different transition types.

4.1. Core concepts

The central element of the proposed re-conceptualization is the comparison target which will be called a character. The size of a character is defined by its attributes. Finally, the context is one of the factors that can make the relationship between characters more complex (compare to section 2.2.1).

4.1.1. Characters char

A character is what we have called a *target* (sec. 2.2.1). In its simplest form, it takes up one slot in the readers' working memory. Characters are clearly distinguishable and nameable visual entities in the chart. It can be a dot, a line or any other mark. In visualization, these entities are very often identified by one or more independent, categorical variables. For example countries, genders, age groups, etc. Even though this definition might seem abstract at first, it is in most cases very easy to identify characters in narrative visualization. That is because the narrative will often explicitly identify them. But characters can also be identified implicitly and through visual statistics.

This indicates that characters are something that only truly exists in the reader's mind and can be dependent on existing knowledge. One example for this is, that multiple characters can be grouped into a unifying character. Switzerland, Italy, and France could form the group *European countries* while Japan, China, and Korea might be combined into *Asian countries*. Grouping is, therefore, one way to reduce the cognitive load by relying on long term memory.

4.1.2. Attributes attr

Attributes define the shape of a character and therefore its complexity. Our classification uses the term to describe all the other variables encoded in the visualization in the form of position, size, shape, etc. Attributes are showing different aspects of the characters. Country-characters can have population numbers over time. Gender-characters might have differing PISA-test success rates or might marry at different ages.

4.1.3. Context ctx

Finally, many charts typically have a context. The situation shown might be for a certain year, for a certain revision of the PISA-test or according to a certain source. The context is often depicted in the title. When the context changes between two charts, it makes the relationship between the characters more complex.

The three concepts roughly mirror McCloud's [25] and Cohn's [30] *subject*, *scene* and *time*. Based on Cohn's work and the discussion in section 2.2.1 we suspect that there is also a hierarchy in our proposed types: Readers will first identify the *characters* in a chart and look for the corresponding characters in the other chart. If they find them and they visually differ between the states, they will look for reasons. At first, they will assume that a different *attribute* of the character is shown and therefore check the axes. If they conclude that the changing appearance of the character is not due to different attributes being shown they will assume that a change of *context* has happened.

4.2. Case study method

Characters, *attributes*, and *context* represent different things that can change between two charts. The proposed reconceptualization, therefore, classifies transitions according to wether *characters*, *attributes* or the *context* change (and how). To find the different transition types, we have applied a case study method that is similar to the study conducted by Hullman et al. [32] but more limited in scope.

4.2.1. Selection of examples

The test the proposed model, we have applied it to a collection of transition techniques commonly found in narrative visualization. We have started by compiling a corpus of narrative visualizations from online sources. The corpus combines two collections from other authors [14], [66] and our examples. The corpus thus includes 144 narrative visualizations published between 2008 and 2019 by a variety of news organizations as well as individuals. The complete corpus can be found in appendix B.

In a first step, we excluded examples that were either not focused on data visualization or examples that did not contain any transitions. We also excluded videos and everything that was done in flash for practical purposes. From the remaining 79 examples we selected 20 examples that we felt were representative of the state of the art based on the following criteria: We preferred more recent examples to older ones. We preferred examples that were heavier on the visuals. And we included a variety of sources.

4.2.2. Analysis

For each example, a screenshot of each chart was pasted on a canvas in order. First, we identified the characters in the first screenshot. For each transition, we annotated if the characters, scene or context changed from the previous state and how it changed. This was usually straightforward which was an encouraging sign. Finally, we unified very similar transitions into 9 categories as presented in table 2. The complete analysis can be found in appendix C.

4.3. Transition types identified

Table 2: Transition types are defined by how characters, attributes, and context differ between two charts.

	Characters char	Attributes attr	Context ctx
Explore attributes	Same	Differ	Same
Contrast characters	Differ	Same	Same
Reconfigure	Same	Get remapped	Same
Split characters	Split	Same	Same
Merge characters	Merged	Same	Same
Highlighting	Less	Same	Same
Progressive disclosure	More	Same	Same
Context	Can differ	Same	Differs
Semantic field	Differ	Differ	Can differ

4.3.1. Explore attributes

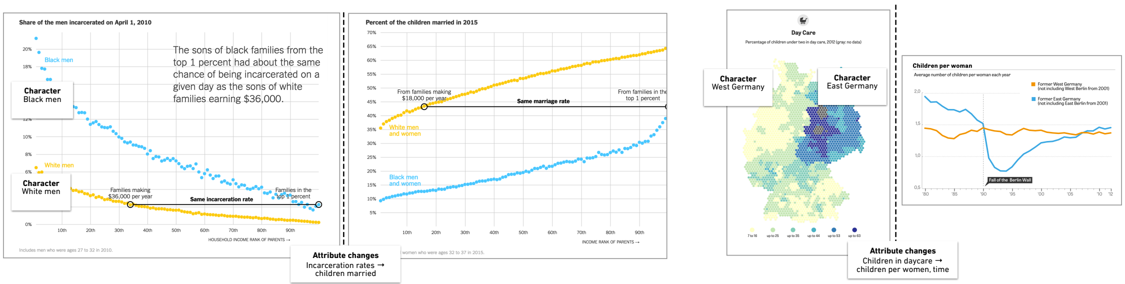


Figure 11: Two examples of transitions exploring different attributes of the same characters. On the left side from a New York Times story [67] and on the right side from a story by "Die Zeit"[10].

Characters char	Attributes attr	Context ctx
Stay the same	Differ	Same

This is probably the most obvious transition type for storytelling. Here the author explores different aspects of the same character(s). In many cases, this will be done by mapping a different attribute to one of the axes, like on the left side of figure 11. Here the author compares incarceration rates attr for black char and white men char by the income of their parents attr in the first chart. In the second, the incarceration rate attr gets replaced with the percentage of children who are married attr . To show that you are more likely to get incarcerated and less likely to get married when you are from a black family.

Note how the character changes technically between the two states as it is first "black men" and after that "black men and women" . This will likely get unnoticed by a majority of the readers by design. Choosing the same colors for semantically very similar character shows that the author intended them to be perceived as "the same".

The example on the right side of figure 11 shows two interesting things: characters need not be explicit and characters can stay the same while chart types differ.

In the first chart, the characters are not identified by the author but emerge from the visualization through a combination of *identification through visual statistics* and *implicit identification* (sec. 2.2.1). More specifically the reader will perceive two clusters of very different color on the map. If he has some knowledge about the geography and history of Germany, he will *implicitly* identify them as Former East char and West Germany char (see section 2.2.1).

These characters are named in the next state. In contrast to the example on the left, the second chart here is completely different from the first. While the first depicts a map, the second is a time-series chart. Nonetheless, the transition is still between two characters showing different attributes (daycare attr and children per women, time attr) for the same characters.

The corresponding transition in Hullman et al. is a *measure walk* (sec. 2.1.3).

4.3.2. Contrast characters

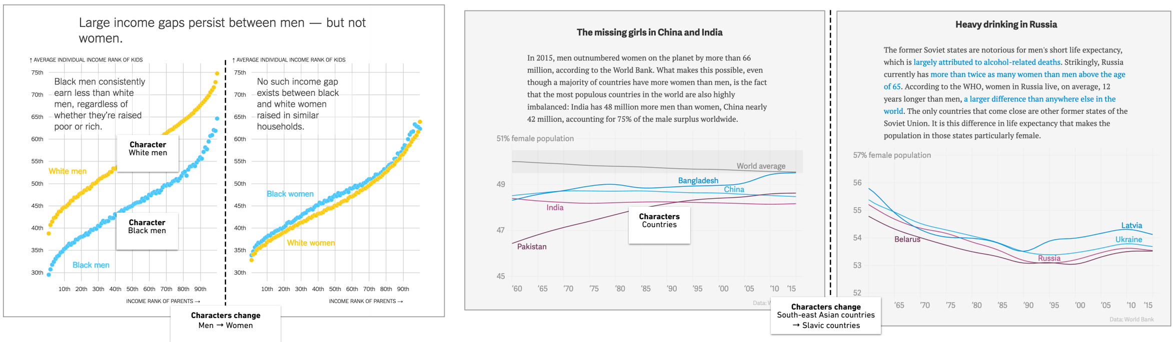


Figure 12: Two examples of how characters are being contrasted. One from a Quartz article 68 and the other from the New York Times [67].

Characters char	Attributes attr	Context ctx
Differ	Stay the same	Same

We call this technique *contrasting* because it focuses on the difference between characters. It's typical for these transitions to keep the same layout between two states and just switch characters. Only like this, a visual comparison is possible. This is the case for the example on the left in figure 12. The income gap between black char and white men char becomes very clear in comparison to the obvious lack of the gap for women char when plotted in the same coordinate system.

But also in contrasting transitions, there is subtlety as visible on the right side of figure 12. Here the x-axis is being shifted between the states. But the shift has no other reason than saving space. It only takes slightly more effort to understand that the first chart shows states that have below-average proportions attr of women and the second chart shows states that have above-average proportions attr of women. But this *pan* (see section 2.1.1) could have been left out without changing the interpretation of the transition at all. This leads us directly to the next technique: *Reconfigure*.

The corresponding transition in Hullman et al. is a *dimension walk* (sec. 2.1.3).

4.3.3. Reconfigure

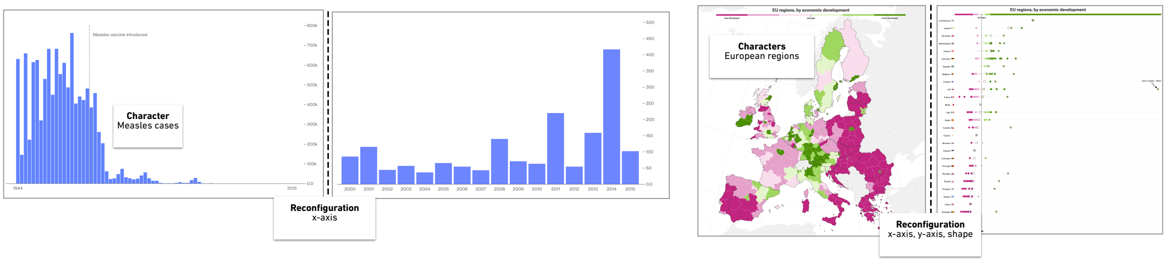


Figure 13: Two examples of *reconfigure* transitions. On the left a *zoom* from a Bloomberg article [69] and on the right a remapping of attributes to different visual variables found in an article by The Pudding [34].

Characters char	Attributes attr	Context ctx
Stay the same	Stay the same but get remapped to different visual variables	Same

The defining characteristic of this transition is, that no new information is shown from the first to the second state. The information is just shown in a different way to clarify different aspects. The left example in figure 13 first shows just how much the cases of measles have been reduced through vaccination. It then *zooms* in on the portion of the chart showing the cases in recent times. They were too small to be visible in the first chart but technically present. This sort of transition is a good replacement for a log-transformed axis.

The right example in figure 13 shows a more drastic reconfiguration of a map into a scatterplot. While the first chart makes it easy to see how the south of Italy is underdeveloped and while the north is above average, the same split can be found in the scatterplot when hovering over the dots, representing the individual regions of Italy. Conversely, the ranking and uniformity of countries can be extracted through visual statistics (2.2.1) from the first chart. It's just much, much clearer in the second.

Reconfigure subsumes the *pan*, *zoom*, *reorder* and *visualization change* transitions from visualization-oriented transitions (see section 2.1.1). We have not found reconfigure transitions to be common enough in narrative visualization to deserve further specification into these subcategories.

4.3.4. Split characters

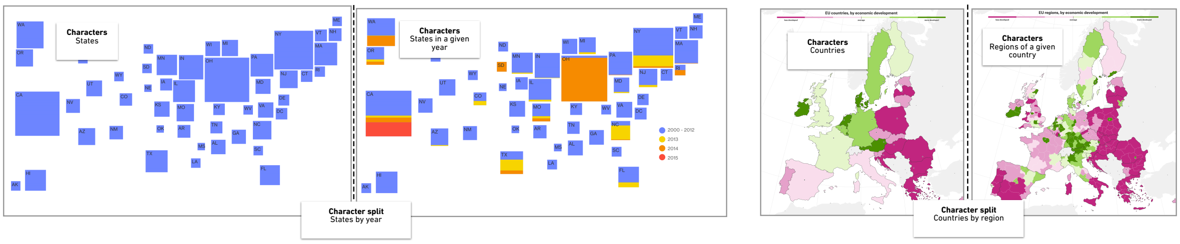


Figure 14: Examples of transitions where characters are being split. On the left measles cases by year [69] and on the right EU countries by regions [34].

Characters char	Attributes attr	Context ctx
Visible characters are split into sub-characters	Stay the same	Same

As described in section 4.1, characters need not be atomic units. Groups of similar characters can be perceived as a character themselves. But sometimes the author wants to convey how characters in a group differ from each other and thus how (in-)homogeneous a group is.

Splitting characters is an essential device in narrative visualization because it nicely fits the general-to-specific pattern often used in storytelling [32]. When splitting characters, the attributes will usually stay the same to serve as a frame of reference. A kind of shadow or contour of the “parent”-group will often remain too for reference. This is the case for both examples in figure 14 one who splits measles cases by states char and years attr and one who splits European countries char into regions char .

Splitting characters is what Hullman et al. call a *general-to-specific* transition (see section 2.1.3).

4.3.5. Merge characters

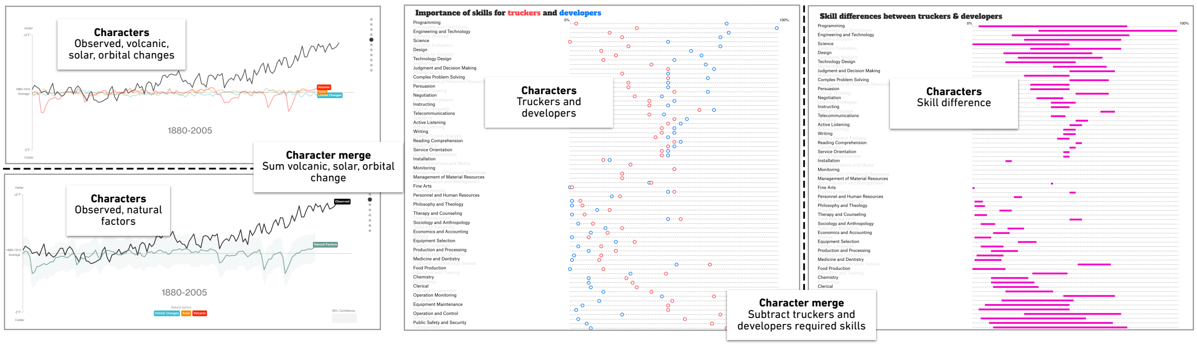


Figure 15: Different ways how characters are merged. On the left a transition where characters are "summed" from from Bloomberg [33] and on the right a transition where characters are "subtracted" from *The pudding* [9].

Characters char	Attributes attr	Context ctx
Visible characters are merged	Stay the same	Same

Where there is splitting, there needs to be merging. But merging is conceptually more complicated than splitting. Splitting shows more information in the second chart while merging summarizes information from the first chart. In the second case, it is much more important for the reader to understand the summary operation that is being used. Is the new character the *mean*, the *sum* or even the *difference* of the characters previously seen? Both examples in figure 15 illustrate this problem very well. Visually, the natural factors char presented in the second chart could very well be the *mean* of volcanic char, solar char and orbital change char. But it's actually the *sum*. In the example on the right, the pink bars might well represent the *sum* of the dots on the left but it's actually the difference. Although merging is complex it can be very useful for a storytelling approach that has been called *ladder of abstraction*-storytelling [70]. It starts with very concrete, down to earth characters (like volcanoes char or skill importance char) and moves "up" towards more abstract concepts (like natural factors char or skill importance difference char). Previous authors have had some success with depicting such operations through animation [53].

Hullman et al. call this transition *specific to general* (see section 2.1.3).

4.3.6. Highlighting

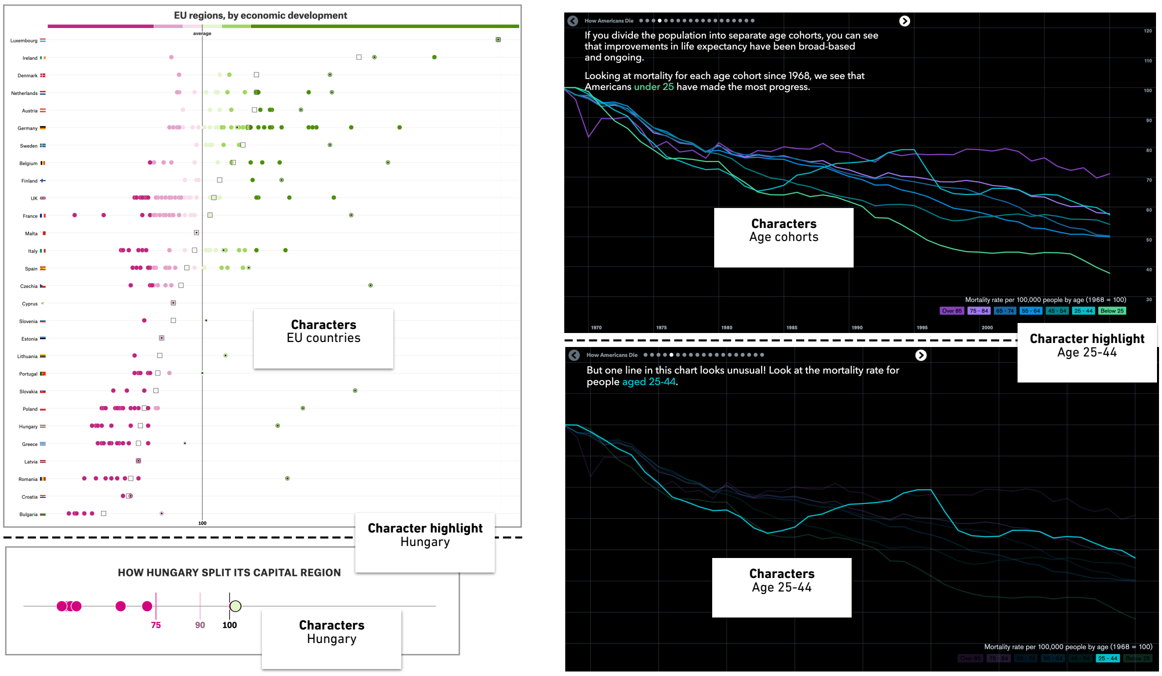


Figure 16: Two examples of *highlighting* transitions. On the left from The Pudding [34] and on the right from Bloomberg [71].

Characters char	Attributes attr	Context ctx
Get filtered down	Stay the same	Same

This technique again serves the very common *general-to-specific*-pattern [32]). It focuses the reader on a specific character or a special group of characters. The most common way to do this is by hiding the other characters. An example of this can be found on the left in figure 16 where the second chart only shows Hungary char from all the EU countries char. In the example on the right side of figure 16 the transition hides the other age groups char to highlight the 25 to 44 group char. Many other ways of highlighting one character are imaginable.

Highlighting is the equivalent of a *filter* transition from the visualization-oriented transitions (see section 2.1.1).

4.3.7. Progressive disclosure

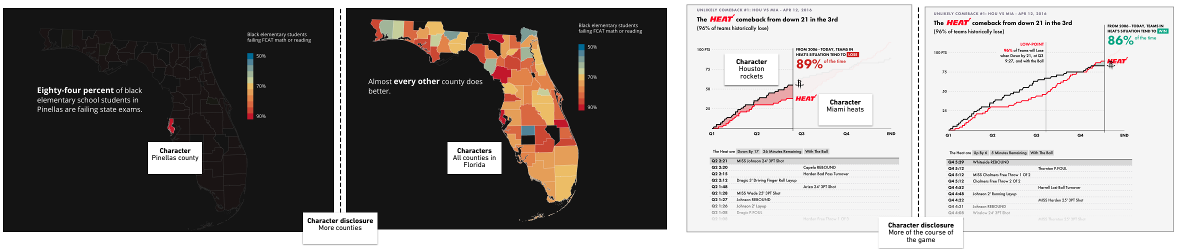


Figure 17: On the left an example of *progressive disclosure* by showing more characters from the Tampa Bay Times [35]. On the right an example where more of the same character is disclosed from The Pudding [72].

Characters char	Attributes attr	Context ctx
More or more of them becomes visible	Stay the same	Same

Similar to *split* and *merge*, this is the counterpart of *highlighting*. And similar, to *merge* it is a useful device for *ladder of abstraction*-storytelling [70]. One version of this technique introduces new characters in the second chart as seen on the left in figure 17. Another variant shows more of the same character(s) in the second chart. The example on the right shows more and more of the course of the game between the Miami Heats char and the Houston Rockets char with each transition.

A particular case of progressive disclosure is what Hullman et al. call *spatial transitions*. A story might, for example, explore the mountainous terrain of the Alps and disclose more and more of it with each transition (see section 2.1.3). But it can also represent the removal of a filter according to the visualization-oriented classification (see section 2.1.1).

4.3.8. Context

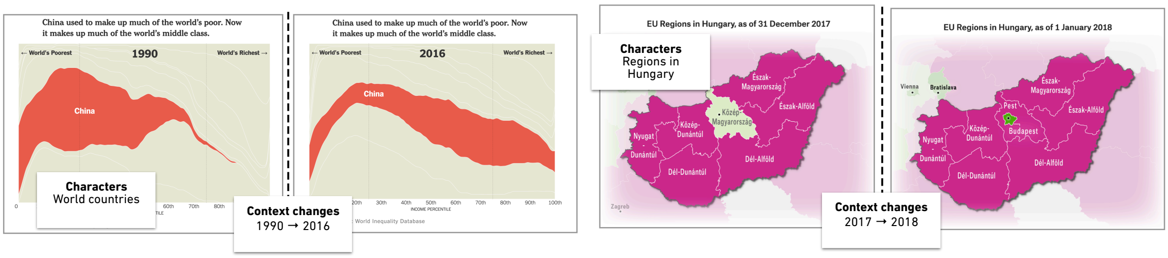


Figure 18: Two examples of transitions where the context changes. On the left from the New York Times [73] and on the right from The Pudding [34].

Characters char	Attributes attr	Context ctx
Stay the same*	Stay the same	Changes

The interest of the context transition is to show how the "situation" (the form of characters) changes when the context changes. The most common case is comparing the situation at different points in time. In the left example in figure 18, the author wants to convey how China char made up most of the world's poor attr in the 90s ctx and is now ctx home to a large part of the worlds middle class attr .

The example on the right shows how a changing context may lead to a change in characters. Here the formerly ctx single large administrative region around Budapest char has now ctx split into the rich Budapest char and the poor Pest char (notice how poor Pest only gets half of the name).

Hullman et al. identified the very prevalent change in temporal context but did not generalize it (see section 2.1.3). As contexts contain groups of characters, they can become characters themselves at certain points of the story. That is why *context* transitions have a close relationship to *contrasting character* transitions and the distinction is not always 100% clear.

4.3.9. Semantic field

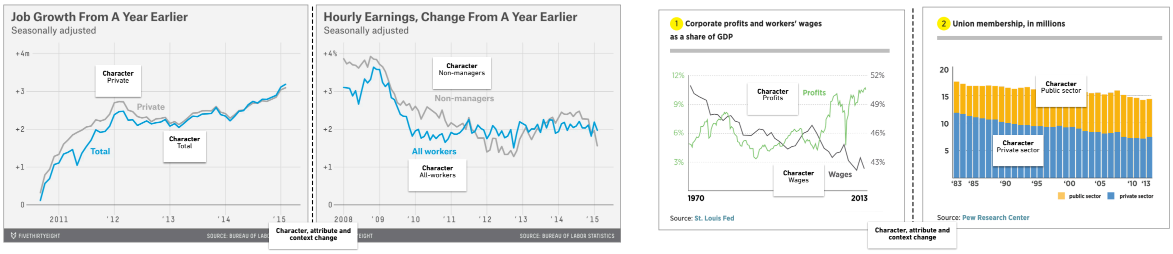


Figure 19: Two examples where charts are only related by a common theme. On the left from FiveThirtyEight [74] and on the right from Vox [75].

Characters char	Attributes attr	Context
Differ	Differ	Can differ

This might be the hardest transition for readers to follow as none of the elements are shared between the charts. The only thing that connects the two charts is that their characters come from the same field of concepts. The reader, therefore, needs to move up in the conceptual hierarchy to find the relationship between the characters depicted. In the example on the left of figure 19, the reader needs to understand how job growth attr is related to hourly earnings attr . This requires a rather sophisticated internal model of economics, or a textual explanation. The same is true for the example on the right. It is not possible without a lot of additional knowledge about how the rising discrepancy of profits char and wages char is related to the decline in union memberships attr .

Causal and *Question & answer* transitions identified by Hullman et al. would fall into that category (see section 2.1.3).

4.3.10. Discussion

The case studies in this section demonstrate how *characters*, *attributes*, and *context* can be used to express a wide range of transitions in very different narrative visualization. It also shows that, even when transitions are looked at through this lens, similar transition types were identified as in Hullman et al.'s study (sec. 2.1.3).

Because *characters* link the marks that are present in the visualization with how the reader interprets them, they introduce the factor of reader knowledge (sec. 2.2.1) to the classifications typically used (sec. 2.1.1) in visualization. We believe that this factor is essential in the context of narrative visualization.

5. Animation based on transition type

5.1. Reducing the transition cost

One of the goals of using animation to guide the reader through narrative visualization should be to *reduce the transition cost*. This can either be done by reducing the number of targets that need to be compared, the size of the targets or the complexity of relationships (sec. 2.3). This section will discuss how the strengths of animation (sec. 3.1) might help to reduce the transition cost. It will then show how different transition types might benefit from different strengths of animation.

5.1.1. Object constancy

Object constancy should make it easier to locate the corresponding target in the second chart. Fluent animation should be a way to offload the task of identifying two objects as "the same" from working memory to the visual system. This, in turn, should considerably reduce cognitive load.

5.1.2. Explicitly show complex relationships

Complex relationships take multiple steps from the reader to understand through comparison. The intermediate results from these steps need to be stored in working memory. Animation might be used to indicate the type of relationship. It might, for example, be used to indicate characters are grouped in a merge transition (sec. 4.3.5).

5.1.3. Highlighting

Highlighting explicitly identifies the targets that the reader should compare (see section 2.2.1). It therefore also reduces the number of targets that need to be held in working memory. Highlighting can be done through the means of narrative visualization (narration, annotations, color, and animation) mentioned in section 1.1. When highlighting is done through pre-attentive attributes like color, for example, objects of different color are already being hidden in the visual system and do not even reach working memory [76] (compare also "Visual feature" in figure). Something similar might happen when using animation to highlight certain targets.

5.1.4. Engagement

While higher engagement might not directly lead to a reduction of the transition cost, it might motivate the reader to focus on the story and therefore still lead to higher rates of understanding. As discussed in section 3.2.1, higher engagement has also been found to be related to congruence.

5.2. Implications for transition types

The transition types presented in section 4.3 will be interpreted in different ways by the readers. While one type will require the reader to find out how the shape of a character has changed, another will require of him to identify characters that have similar shapes, and a third one will require of him to create a relation at the conceptual level. Not all of them will, therefore, benefit from animation in the same way (sec. 5.1). Table 12, therefore, links the strengths of animation to the different transition types.

Table 12: How different transition types can benefit from animation.

	Object constancy	Explicitly show relationship	Highlighting	Engagement
Explore attributes	✓		✓	✓
Contrast characters			✓	
Reconfigure	✓			
Split characters		✓	✓	✓
Merge characters		✓	✓	✓
Highlighting			✓	
Progressive disclosure			✓	
Context	✓		✓	✓
Semantic field			✓	

Object constancy Requires that characters are shared between charts but that they change their location or form. This is because locating the corresponding character in the second chart is only difficult if it is visually different from the first chart. This is true for *explore attributes*, *reconfigure* and *context* transitions.

Explicitly show complex relationships When the relationship between the characters in the two charts is not one-to-one, the reader may benefit from explicitly showing the relationship through animation. This is true for *split characters* and *merge characters* transitions.

Highlighting Different f§rom the other two benefits, highlighting does not require characters to be shared between the states. All transitions can therefore potentially benefit from highlighting. The exception are *reconfigure* transitions because they typically concern all the characters. No highlighting needed.

Engagement Because in previous studies, only congruent animation has been shown to lead to higher engagement. Animation for higher engagement would, therefore, presume some kind of natural motion between characters. This excludes *contrast characters* and *semantic field* transitions. Animation for *progressive disclosure* also has not led to higher reader engagement in a previous study [54]. *Reconfigure* transitions are by nature very abstract and animation would, therefore, most probably not lead to higher engagement.

6. Research hypotheses

Based on the discussion in the previous section (sec. 5.2), we believe that animated transitions have the potential to support readers in understanding the relationship between two charts. This is the main research question of this work as mentioned in the introduction (sec. 1). H2 to H6 test more specific hypotheses based on the strengths of animation presented in section 3. The hypotheses only apply to transition types where animation can theoretically provide a benefit (compare to section 5.2).

H1: Animation makes understanding transitions easier Based on the widespread use of animation in practice and based on the discussion in this chapter, we hypothesize that animation would reduce the cognitive load of readers. Because they would have more available working memory capacity, we hypothesized that they would more likely interpret a transition correctly.

H2: Animation implies a relationship Based on the findings that animation can communicate relationships (see section 3.1.3), we hypothesized that animated transitions might imply a relationship in a more general sense. We hypothesized that animated transitions would more often lead participants to assume a relationship between two charts.

H3: Animation implies a causal relationship Based on the finding that animation can imply causality (sec. 3.1.3) we proposed this as an extension to H2. We hypothesized, that when readers assumed a relationship between two charts they would often specifically assume a cause and effect relationship when they saw an animated transition.

H4: Animation leads to object constancy We hypothesized that animated transitions would support the reader through object constancy.

H5: Animation highlights characters The discussion in section 3.1.2 has shown that the literature agrees that animation would have a highlighting effect. We, therefore, assumed that this would hold for animated transitions in narrative visualizations. Our hypothesis was that animation would focus the reader on the characters that were being animated.

H6: Animation leads to higher engagement Even if it might not directly contribute to transition understanding, we hypothesized that animated transition would lead to higher reader engagement. This hypothesis is mostly based on anecdotal evidence, as discussed in section 3.1.4.

Part III.

Experiment

7. Research method

To test the hypotheses presented in the previous chapter, we have conducted an experiment which will be presented in this chapter. Much this design is based on discussions with Dr. Nicole Jardine, Prof. Dr. Steven Franconeri and Cindy Xiong from the Visual Thinking Lab at Northwestern University, Chicago [77].

7.1. Overview

In the experiment, we presented two narrative visualizations (stories) to each participant: *mortality* and *energy*. Each story contained a total of five charts and four transitions. Each participant saw one of the stories with *animated* transitions and the other with *static* (non-animated) transitions.

To examine each transition individually, we have split each story into four separate screens which we called *mini-stories* (see figure 20). Each mini-story contained one transition and the two corresponding charts.

This setup has made it so that the second chart of each mini-story is the first chart of the next mini-story. Normally, this would not be true for the initial mini-story. Because we did not want the initial mini-story to differ from the others in this respect, we have introduced another screen before each mini-story where the first chart is presented (see figure 20)

At the end of each mini-story, we presented a questionnaire asking about the participant's conclusion and their perceived engagement. The questionnaire will be presented in more detail in section 7.4.1.

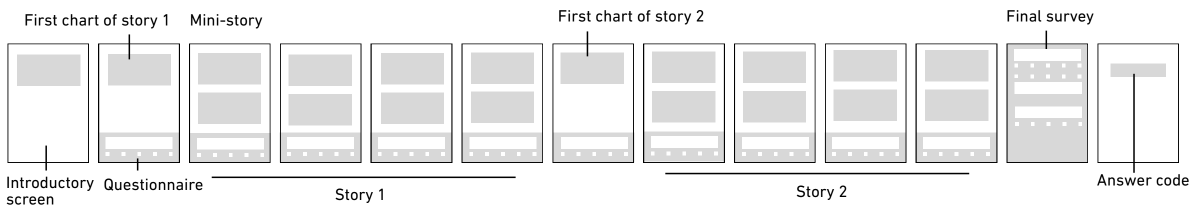


Figure 20: The order of the 13 screens presented to the participants in the experiment.

The experiment was conducted on Amazon Mechanical Turk [78] and had 56 participants. An introductory screen instructed the participants on what they would have to do and that they would receive

USD 9 for completing the experiment. After the two stories, a final survey (sec. 7.4.2) and finally an answer code to verify their participation was presented. We also limited participation to U.S.-citizens because the experiment was in English and concerned topics that treat the U.S.

7.2. Stimuli

This section will present the stimuli, the *transitions* and the *stories* in more detail. This will make the presentation of the experiment in the following sections much clearer.

7.2.1. Transitions

Each participant saw both stories and each story contained either *animated* or *static* (non-animated) transitions. This design ensured that the distinction between *static* and *animated* transitions had maximal statistical power.

In both cases, the transitions were controlled by scrolling. Initially, a participant would see the first chart of a mini-story. When he scrolled down, the first chart would be replaced by the second chart. When scrolling even further, the questionnaire would appear. In *animated* transitions, the chart would not simply be replaced but morphed into the other. The animation was fully controlled through scrolling and could be played forward and backward. This ensured that the amount of *direct manipulation* would be the same for both transitions (sec. 3.2.3). The animation was also implemented with “slow-in-slow-out”-easing which has been shown to work better than linear easing (sec. 3.2.3).

7.2.2. Mortality story

The first story was about the evolution of mortality rates in the U.S. It is a modification of an article published by Bloomberg in 2014 [71]. The individual charts are displayed in figure 21.

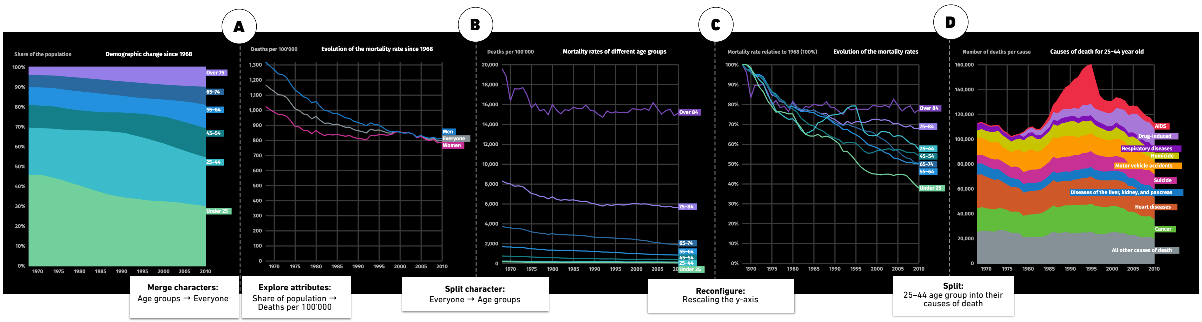


Figure 21: The charts and the transitions of the mortality story.

Mortality A The story begins with a chart showing how the distribution of age groups char has changed resulting in an aging population char over time attr . The transition first *merges* the individual age groups char into an "Everyone" group char . Then it replaces the share of population attr with the "deaths per 100'000" attr . The next chart shows how "deaths per 100'000" have been reduced drastically over time attr . The transition thus conveys how reducing mortality rate leads to an aging population.

Mortality B The second transition *splits* "Everyone" char again into the different age groups char . From this mini-story, the reader learns that the mortality rate is dominated (unsurprisingly) by elderly people. An apparent slowdown in mortality rate reduction in the first chart is therefore mostly due to to the development of the "Over 84" group char .

Mortality C The third transition *reconfigures* the x-axis to show the relative evolution of the mortality rate attr for each age group char . Thanks to this it becomes apparent that even though the elderly have made the most progress in absolute numbers, the relative reduction is highest for younger age groups.

Mortality D Some readers might have wondered why the mortality rate for "25–44" year olds char was rising at the beginning of the 90s. The fourth transition, therefore, *splits* this age group char by its causes of death char . This answers the question: it was precisely at this point that the AIDS-epidemic was at its peak and has most strongly affected this age group char .

7.2.3. Energy story

This is a story about the evolution of energy sources in the U.S. since the beginning of the century. It combines a story from the New York Times [79] with one from Forbes [80]. The individual charts are presented in figure 22.

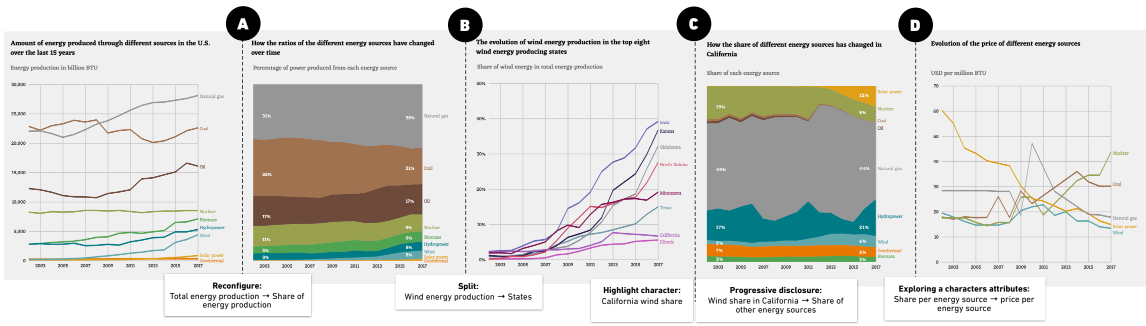


Figure 22: The charts and the transitions of the energy story.

Energy A The energy story starts by showing the evolution of different energy sources char in the U.S in absolute numbers attr . The transition *reconfigures* these characters to show how the proportion of use attr for each energy source char has changed over time attr . From comparing the two charts it becomes clear that while the absolute output has either risen or stagnated, coal char has been used significantly less while the focus seems to be shifting to natural gas char and the renewables char . The animated transition morphs the lines into their respective areas in the second chart.

Energy B The second transition *splits* wind energy char into different states char . It shows how wind has become a major energy source in certain (great plains) states and leads to the conclusion that, although the rise of wind energy looks unspectacular on a national scale, growth has been exponential in some places.

Energy C The third mini-story focuses on the strange stagnation of wind energy ctx in California char after 2013. The third transition first *highlights* wind energy in California char and then *discloses* the usage proportions attr of other energy sources char in the same state ctx . A marked rise in solar power char after 2013 implies that the state has decided to invest in this renewable instead of wind char .

Energy D The final transition *explores* one of the possible reasons for Californias ctx decision to focus on solar energy from 2013 on. Instead of the shares attr of the individual energy sources char , it shows their cost attr . While wind energy char has been consistently inexpensive, the price of solar power char has fallen substantially over time. It even reached the price level of wind energy in 2013. Which implies that Californias focus on solar rather than wind energy might have economic reasons.

The presented stories were implemented with web technologies and viewable in the browser. A more in-depth discussion of the implementation can be found in chapter 11. The online versions of the stories are referenced in appendix D.

7.3. Conditions

The main goal of the experiment was to find differences between *animated* and *static* transitions. We, therefore, chose this to be the *independent variable*. We counterbalanced for some of the confounders like *topic* and visualization *design* by using the two different stories with the different designs presented before. This has resulted in the following within-subject design:

Condition	Story 1	Story 2
1	Mortality static	Energy animated

Condition	Story 1	Story 2
2	Mortality animated	Energy static
3	Energy static	Mortality animated
4	Energy animated	Mortality static

7.4. Data collection

7.4.1. Questionnaire

Conclusion The questionnaire at the end of each mini-story asked participants the following question: "What is your overall conclusion from the two charts on this page?". It then provided a free-form text field for them to provide their answer. We asked for answers that were at least three sentences long. The wording was chosen so as to not hint at a relation between the two charts. We have included two other questions that served as attention and understanding checks. The full questionnaire can be found in appendix E.

Focused attention The questionnaire also included a subset of the validated *user engagement scale* questionnaire. The scale measures multiple attributes like *focused attention*, *perceived usability*, *aesthetics* or *reward* that together form what is known as *engagement* in the literature [81]. But even the short questionnaire contains twelve questions which would have been too long to ask after each transition. We, therefore, decided to use a sub-scale of the *user engagement scale* that measures *focused attention*. The reason to use *focused attention* was that it was found to be the factor that explained most of the variance in O'Brien et al.'s analysis [81]. Also, the questions skew towards the positive (compared to the other option, *perceived usability*) and we did not want to negatively prime the participants.

7.4.2. Survey

The survey at the end of the experiment was introduced to get a sense of the demographics and some additional information that might help to explain outliers.

Age In the final survey, we asked the participants about their age to better understand the diversity of the population that was participating in the experiment.

Gender We equally asked the participants about their gender to help us understand how diverse the population was.

Level of education This would act as a rough proxy for visualization understanding. We assumed certain outliers might potentially be explained by a low level of education.

Vision Equally, if the participants reported weak and uncorrected vision, we could potentially exclude them from the analysis if we found that the factor skewed the results.

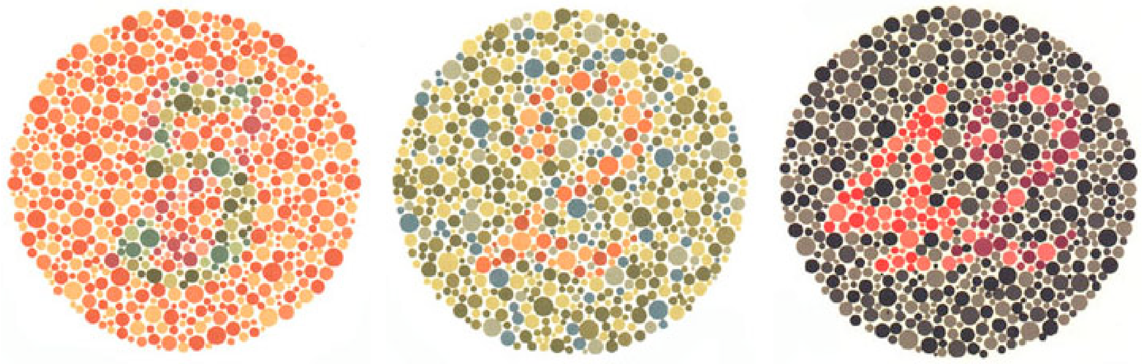


Figure 23: The three plates of the Ishihara colorblindness test [82] that were presented to participant in the final survey.

Colorblindness A final factor we considered for explaining outliers was color blindness. We did not ask participants to report on their colorblindness because not all people who are colorblind know about it. Instead, we presented them the three plates from the Ishihara colorblindness test shown in figure 23. The plates we have selected, tested for red-green blindness. The most common form of colorblindness, that affects about 8% of men and 0.4% of women [83].

The full survey can be found in appendix F.

7.4.3. Passive data collection

The participants actions on each page were continually monitored. We recorded scrolling, typing, and information about the participant's browser throughout the experiment.

Scrolling Each scrolling interaction with a transition was recorded and timestamped so that we would be able to review how many times a participant saw a transition and how they played the animated transitions.

Typing Each keystroke was recorded with a timestamp. This allowed us to separate viewing time from typing time.

Start and end times Were recorded to know how much time each participant spent on each mini-story.

Browser parameters Browser parameters like window size and user agent were collected as another way to explain outliers.

A complete overview of the data that was collected passively can be found in appendix G.

7.5. Measures

To test the hypotheses (sec. 6) and based on the data collected (sec. 7.4) we have established several measures.

H1: Transitions understanding The main hypothesis was that the combined benefits of animated transitions would support readers in understanding the relationships between charts. We, therefore, coded their conclusions (sec. 7.4.1) based on the relationship that was implied by each transition (sec. 7.2.2 and sec. 7.2.3). This led to a measure called **correct relationship**.

H2: Relationship In addition to coding the *correct relationship*, we also coded if participants assumed any kind of relationship between the two charts in their answer. This measure was called **relationship**.

H3: Causality Whenever the participant perceived a relationship, we also coded if his answer indicated that he assumed that the relationship was causal. We called this measure **causal**.

H4: Object constancy If animated transitions would support object constancy, participants would need fewer transitions to come to their conclusion because their cognitive load would be reduced. We, therefore, measured the **number of transitions** between the two charts. When the transition was animated the participants needed to cross the 10% and then the 90% threshold successively (or the inverse) for it to be counted as one transition.

We presumed that because object constancy would reduce the number of comparisons needed, animation would also lead to a lower **viewing duration**.

H5: Highlighting If animation has a cueing effect, it should focus the participant on certain characters. We predicted therefore that this will lead people to mention fewer characters in their conclusions. The **number of characters mentioned** in each answer was therefore coded. We not only counted characters that were directly visible but also groupings of characters like "green energy sources" (sec. 4.1).

H6: Engagement Engagement was measured by calculating the mean of the answers on the **focused attention** sub-scale of the *user engagement questionnaire* [81].

We considered **character count** to be another proxy for engagement. As the participants were in a task-oriented setting, we generally expected them to not lose their time with long answers. If answers would turn out to be considerably longer in one of the scenarios, we assumed that it was because it somehow engaged them.

Finally, we have used *sentiment analysis* on the participant's answers. We assumed, that if there was a difference in engagement, this might lead to a more confident or positive tone in the answers. To analyze the answers, we have used IBM Watson [84] which gives ratings for different types of tones in the text. Watson identifies the degree to which a text is *analytical*, *confident*, *tentative*, *fearful*, *joyful* or *sad*.

Because we would not expect highly emotional language in the participant's answers, we decided to look further only into the three dimensions: **analytical tone**, **confident tone**, and **tentative tone**.

7.6. Data analysis

7.6.1. Calculation of measures

Total duration The duration was calculated by excluding times of inactivity. We considered timespans of more than 30 seconds without scrolling or typing to be *inactive time*.

Typing time Typing time was calculated by summing all the times between recorded keystrokes (excluding *inactive time*).

Viewing time This was defined as the *total duration* minus *typing* and *inactive time*.

Scroll position The scroll positions were normalized to the distance between the beginning of the first chart and the beginning of the second chart. This resulted in values between 0 and 1 indicating the progress in the transition.

Transition count Transition count was calculated based on *scroll position* and *time*. Each time the scroll position passed from lower than 0.1 to greater than 0.9 (or the inverse) subsequently, was counted as one transition.

Drawing performance Each time a chart was redrawn during an animated transition was recorded. Because redrawing was coupled to scrolling, it was not always clear if long gaps indicated that the drawing performance was bad or if it was just the participant who scrolled slowly. A threshold was therefore set at 0.1 seconds. Only redraws that happened within this timespan were considered for calculating the mean drawing performance for each participant and mini-story. Everything below 20 frames per second was considered *slow*.

7.6.2. Statistical methods

To evaluate how the conditions differed, the *means* of the above measures where compared. To find if the conditions differed significantly a *two-sided t-test* was employed to calculate the *p-value*. Finally, the *means* with their respective *95% confidence intervals* were plotted against each other to give a clearer picture of the distributions.

The complete analysis can be found in appendix H.

7.7. Design choices

While designing the narrative visualizations a lot of small choices needed to be made. Many of them were based on informal testing and observation of test subjects. This section will give brief argumentations for some of the most important decisions. A longer discussion can be found in appendix J.

Minimal textual narrative While narrative visualization is usually coupled with narrative text (sec. 1), we have excluded longer texts as a potential confounder from the experiment.

Story selection Because text was excluded, the stories needed to be very visual to be understandable solely based on labels and the title. Also, the stories needed to be about topics that the participants knew something about. This is because contextual knowledge is important to understand transitions (sec. 2.2.1).

Chart types We have chosen very simple time series charts because participants should ideally not fail in understanding the basic charts.

Transitions Transitions were designed so that they would ideally *highlight* the characters that showed the relationship to with the second chart. They were then morphed into the related characters in the second chart to support *object constancy*. Other than that they did not contain any information to avoid *unfair comparisons* (sec. 3.3).

7.8. Other factors considered

Screen size To exclude effects that would result from differences in layout and element size, the experiment was designed for a fixed viewport size of 1280px × 720px. Participants with viewports that were smaller than this were blocked from the experiment.

Browser capability Equally, participants with browsers that did not support the technologies we used for the animated transitions were blocked from the experiment. This had the positive side-effect of equally blocking slower browsers.

8. Results

173 participants started the experiment and saw the initial explanations. From these 56 provided a final answer code to Amazon Mechanical Turk. But 3 of these participants did not finish the experiment and provided a wrong answer code. They were therefore excluded from the analysis. This leads to a slightly uneven distribution of participants between conditions. See table 14.

Table 14: Number of participants per condition.

Condition	1	2	3	4
Participants	13	14	14	12

It took participants on average 8 minutes to answer a mini-story. The maximal duration was 15 minutes. Some participants took breaks of considerable length. While the median *inactive time* was 1.8 minutes, the maximum was 89 minutes. Inactive times were therefore excluded from the analysis.

On average, participants took 52 minutes to complete the experiment. This results in a mean hourly wage of USD 10.40 with a minimum of USD 5.56 for the slowest participant.

We have also found that 7 participants experienced animations with a drawing performance of less than 20 frames per second in a total of 16 mini-stories with animated transitions. This corresponds to about 8% of the samples with animated transitions. The mean frame rate was 41 frames per second.

8.1. Demographics of the participants

Figure 24 gives an overview of the demographics of the participants.

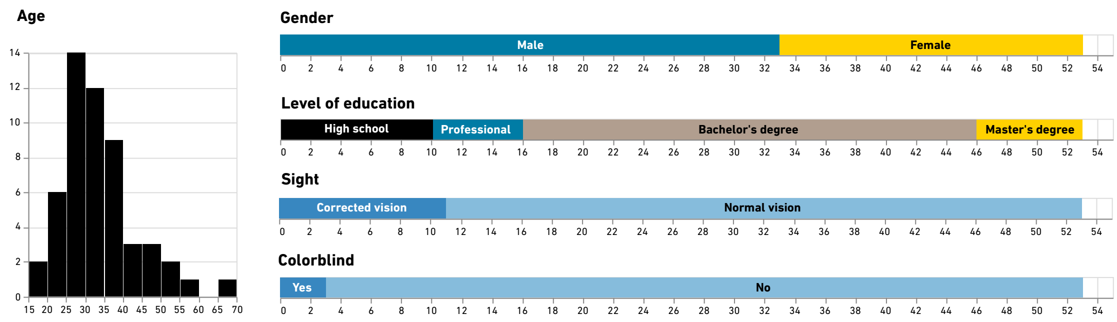


Figure 24: An overview of the demographics of the experiment's participants.

Because the experiment was conducted on Amazon Mechanical Turk, the participants were quite varied though there was a tendency towards men and younger people. 70% of the respondents had a Bachelor's degree or higher.

There were no participants who did not have either normal or corrected to normal vision. The Ishihara plates found three colorblind participants. All of them men.

8.2. Differences between the stories

As discussed in section 7.7, one of the main challenges was to design narrative visualizations that a large part of the participants would understand. Table 15 therefore shows the results for different measures by mini-story.

Table 15: The measures for the mini-stories indicate very different difficulties. In the heading, *e* stands for *energy* and *m* stands for *mortality*.

mini-story	e-A	e-B	e-C	e-D	m-A	m-B	m-C	m-D
correct relationship	19%	24%	19%	47%	33%	14%	12%	37%
relationship	37%	42%	33%	62%	51%	36%	18%	47%
causal	4%	8%	6%	42%	42%	2%	0%	31%
transitions	4.0	4.2	3.4	4.4	3.7	4.7	5.3	3.6

mini-story	e-A	e-B	e-C	e-D	m-A	m-B	m-C	m-D
characters mentioned	2.3	1.7	3.3	2.3	1.0	1.3	1.1	1.5
focused attention	2.9	3.0	3.1	3.0	3.2	3.1	3.1	3.3
answer length (chars)	1417	1322	1203	1244	1244	1205	1226	1275
analytical tone	84%	80%	79%	80%	76%	76%	75%	77%
confident tone	71%	83%	85%	82%	77%	82%	77%	75%
tentative tone	78%	82%	80%	80%	77%	80%	79%	76%
viewing duration	160 s	146 s	124 s	135 s	144 s	132 s	131 s	148 s
typing duration	205 s	200 s	179 s	179 s	212 s	191 s	217 s	214 s

The mini-stories with the highest rate of success (**corrrect relationship** and **relationship**) were *energy D*, *mortality A* and *mortality D*. All of them had also very high rates for **causal** and depicted a cause and effect relationship (sec. 7.2.2 and sec. 7.2.3). *mortality C* had the lowest success rate and also the highest **transition count**. In general, more **characters** were mentioned for the *energy* story which did contain more characters in general (sec. 7.2.3).

There were no clear differences in **focused attention** and **answer length** between the stories. Participants had a slightly more **confident** and **analytical** tone in the *energy* story. In both stories, the **viewing duration** was higher for the first mini-story. There were no clear patterns for the typing durations.

8.3. Hypotheses

We have calculated the *mean, 95% confidence interval* and the *p-value* according to a two-sided t-test for each transition type and measure. The results are displayed in figure 25.

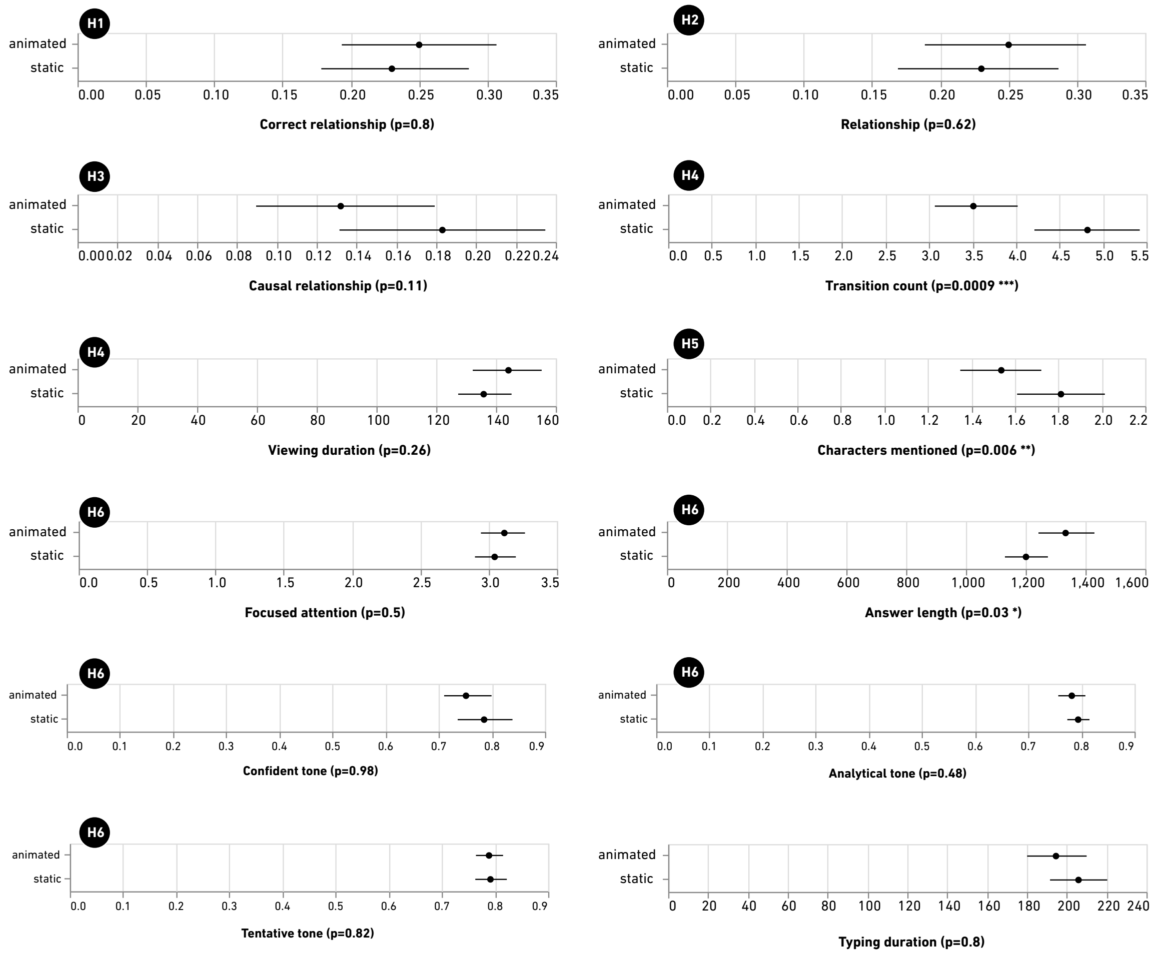


Figure 25: Means, confidence intervals and p-values for the different measures.

H1: Transition understanding We did not find any indication that animation resulted in a better understanding of the relationships that the author wanted to convey. The main hypothesis of the experiment, therefore, needs to be rejected.

H2: Relationship Neither did we find that animated transitions implied a relationship between the charts more often than static transitions. Animation, therefore, does not seem to imply a relationship between two charts.

H3: Causality The results for this measure come close to the 0.05 threshold for significance. But the share of positive samples (causal relationships mentioned) is very low. This means that this result is not very robust.

H4: Object constancy Participants who saw animated transitions did not switch back and forth between the charts as often as those who saw the static transitions. The difference is highly significant.

On the other hand, the *viewing durations* did not differ significantly between animated and static transitions.

Typing duration was included for completeness and did not differ significantly between transition types.

H5: Highlighting The results show that participants who saw static transitions mentioned significantly more characters in their answers.

H6: Engagement Participants did not self-evaluate their *focused attention* significantly higher when they saw animated transitions. Neither did they differ in their tone. But the results for the tone need to be taken with a grain of salt. The Watson API [84] only returns results if the values are higher than 50%. *Confidence* ratings were, therefore, only available for 14%, *analytical* ratings for 69% and *tentativeness* ratings for 37% of the answers.

There was a slight and significant difference between *answer lengths* in favor of animated transitions.

9. Discussion

As presented in section 8, animation does not increase readers *understanding* of transitions between charts. This result is unintuitive for most people because animated transitions "feel" like they help. But our results are consistent with what previous authors found (sec. 3). They found that animation only leads to better understanding when the display is not abstract (as our charts). Even though elements in the chart can be conceptualized as *characters* there is no natural way they should be animated for the types of transitions that we have tested in the experiment. We believe that for similar reasons, animated transitions did not specifically evoke *relationships* between abstract elements in charts.

We think the result that animation does not imply *causality* needs further consideration because it contradicts findings discussed in section 3.1.3. Here it is important to note that only three mini-stories contained causal relationships. This led to a very low number of positive samples and it makes sense that readers would not see causality where there is none. We, therefore, consider the results on causal transitions to be inconclusive.

Readers who saw animated transitions did not need to switch back and forth as often as those who saw static transitions while their *understanding* stayed at the same level. This is a strong indication that animated transitions indeed provide the benefit of *object constancy* as discussed in section 3.1.1. But *viewing times* did not differ significantly. A potential explanation is that while readers transition less, they spend more time on animated transitions moving within the thresholds.

The *highlighting* effect of animation was not disputed in the literature and our experiment also found it. Interestingly, even though it focused readers on the "right characters" it did not lead them to a better understanding of the characters' relationship. This indicates that the difficulty in the transitions did not lie in the number of characters, which generally was within the limits of working memory. The main difficulties were more probably in the *size of the characters* and the *complexity of their relationship* (sec. 2.2.1). Which is good for testing the main hypothesis which was about relationship understanding.

Although many authors are convinced that animation leads to higher *engagement* (sec. 3.1.4) we have found almost no evidence for this. Our results agree therefore with [54] who was the only author in our review who had used a validated scale. Still, these results are counter-intuitive even for the authors of this work. One explanation might be that we have selected the wrong sub-scale. But the lack of a difference in *tone*, even with the reservations discussed, also indicates that the animation did indeed not affect the participants' attitude or mood. A better explanation, therefore, might be that the participants were in a task-oriented setting and the goal of "enjoying" the animation was secondary to them.

An interesting observation is related to the results for individual mini-stories. It seems that transitions that communicate cause and effect relationships are much easier for readers to understand. This is even the case when the transition in itself is rather complex like in *mortality A*. This indicates that transitions are interpreted on an even higher, more conceptual level than we had assumed.

We conclude that while animated transitions facilitate object constancy and highlighting, it is not easy to operationalize these two strengths to support understanding of transitions between charts. We believe that because charts and their relationships are *abstract*, animation is generally not *congruent*(sec. 3.2.1). Therefore, the pre-processing mechanisms that humans have for motion in the real world can not be used to facilitate the interpretation of abstract relationships.

10. Limitations

The main limitation of the proposed experiment is that it did not specifically consider the transition types described in (sec. 4.3). This has led to transitions that combine multiple types like *mortality A* (sec. 7.2.2) and makes it hard to relate the results directly to the concepts that have been established in section 4. It is, therefore, possible that animated transitions are indeed beneficial for very specific types of transitions. We did not test for specific transitions types, because this would have required a more complex experiment for which we lacked the resources. A second reason is, that the theory developed in parallel to the creation of the experiment, which resulted in a lack of time.

A second very important limitation is the high difficulty of the transitions in the experiment. Ideally, each story should have rates of about 60% for the *relationship* measure and around 40% for the *understanding* measure. Only three stories came close to these values. We have also found ordering effects in the results. It is therefore not ensured that all mini-stories are comparable.

A third limitation is the selection of measures and the interpretation of the data. Most hypotheses could only be measured via a proxy variable. Object constancy for example via the transition count. This introduces ambiguity to the interpretation of the results. Additionally, many of the measures were based on a coders interpretation of the answer. At the time of writing most of the answers were only coded by a single person which introduces a lot of subjectivity (although we made sure that the coder did not know which transition an answer corresponded to).

A fourth limitation is the performance of the animated transitions. We did not exclude the 8% of animated transition sessions that were classified as slow in our analysis because this would have strongly skewed that distribution between the groups. But it is likely that lagging transitions had a negative impact on many of our measures.

The results are also limited by the demographics of the participants. They might not apply to readers from other cultural backgrounds and with a different level of education. Also, there were very few old people in the study. The results in a way, therefore, may only apply to the "high performers": young, very well educated Americans.

Finally, although we have counterbalanced the designs, the results may not generalize to other visualization designs than the two time-series charts tested.

Part IV.
Implementation

11. Implementation

To quickly create narrative visualizations and animated transitions for the experiment presented in the previous chapter, we have built a visualization tool that would simplify this task. The implementation builds on the concept of *characters* introduced in section 4.1. The following sections discuss how this contributes a new way of configuring animated transitions in visualization tools and give an overview of the implementation.

11.1. Requirements

This section presents the requirements on the implementation.

Online narrative visualization Much of the specification of the experiment has already been described in chapter 7. To conduct the experiment on Amazon Mechanical Turk [78] we needed a tool to build web-based narrative visualizations with animated transitions. It was, therefore, an obvious choice to use common web technologies like HTML, CSS, SVG, and ECMAScript for the implementation. For increased programmer convenience through type checking, we have used TypeScript to generate the final ECMAScript.

The D3 library [85] was used because it abstracts some of the more tedious parts of generating SVG-code from data. D3 was chosen because it is currently the de-facto standard for data visualization on the web and it does not impose a certain structure on the code as many frameworks do.

Rapid prototyping Because the stories and the design were developed alongside with the implementation, we wanted a system the would allow us to quickly test different design choices or modify the stories. We, therefore, decided to build an application generator the would interpret a domain-specific language (DSL) perfectly tailored to our needs.

Limited range of chart types The range of charts that the general public easily understands is very limited [86]. The tool therefore only needed to be able to generate a limited number of basic chart types.

Transitions focused on characters Based on the proposed concepts presented in chapter 4, we decided to base all of the animated transitions on the concept of *characters*.

Performance Performance needed to be good enough to not impair the perception of the animated transitions. We made some informal user tests to find out what would be an acceptable performance benchmark. For this, we presented two versions of the same narrative visualization to participants (n=4). One which presented the animation at about 15 frames per second (fps), the other at about 40 fps. After much consideration, two of the subject correctly identified the slow animation. We, therefore, gave it some margin and set the lower performance limit to 20 fps which is in line with the 25 fps typically used for movies.

Reliable logging of user interaction and answers Finally, the implementation needed to provide a reliable way to collect the passive data described in section 7.4 and make sure that all the answers were transmitted and that everything could be reliably connected to the right participant.

11.2. Prior art

Many visualization tools have been created over the years and every tool addresses a different set of requirements and has different strengths and weaknesses. This section presents a very brief overview of the state of the art in visualization tools with a focus on the use of domain-specific languages and animated transitions. A more in-depth overview of the state of the art of visualization tools can be found in a very recent paper by Mei et al. [87].

11.2.1. Domain-specific languages for visualization

The number of useful chart types in data visualization is surprisingly limited. A review of some catalogs of chart type shows that it lies somewhere in the range between 40 and 200 [88]–[91]. When developing software for data visualization, an obvious approach is, therefore, to implement these basic chart types and make them configurable. A recent study on visualization tools [87] shows that over the last three decades a majority of tools mentioned in the literature have used some sort of chart typology (see figure 26).

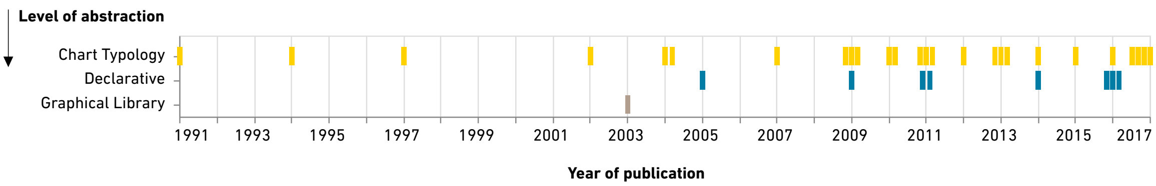


Figure 26: Visualization tools according to their level of abstraction and their year of introduction.

Graphic libraries inhabit the other end of the spectrum. They usually only provide functionality for drawing basic shapes. They are thus the most general and flexible visualization tools. But building visualization out of basic shapes is very laborious, which is the reason why “chart configurators” are so popular. More recently, visualization researchers have started to explore the middle ground: declarative, domain-specific languages (DSLs) [92]–[94]. DSLs try to maintain a maximum of flexibility while significantly simplifying the creation process [93]. One very recent example of a declarative language for visualization on the web is Vega-Lite [94]. Vega-Lite’s DSL builds on JSON and the concept of the *Grammar of Graphics*. The grammar of graphics separates different parts of a visualization into separate “layers” which are independent and can be recombined as desired (see figure 27). Vega-Lite also introduced a way to define interactive visualization in a declarative manner.

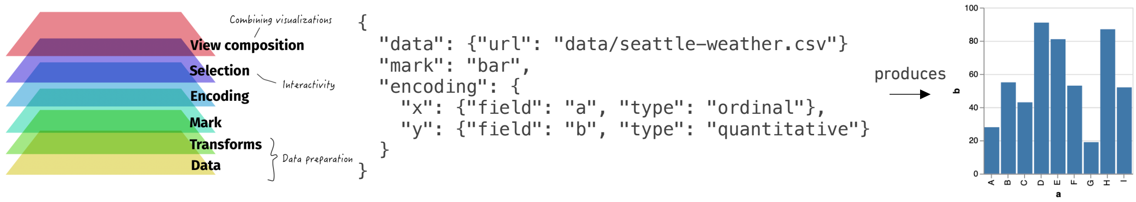


Figure 27: An example of how Vega-Lite representes the different layers of the Grammar of Graphics and how this is translates to a visualization.

11.2.2. Animated transitions between charts

Even less research has been conducted on DSLs who describe animated transitions between visualizations. Heer and Bostock, who later went to create the already mentioned D3 [85], have described one approach in an earlier paper [93]. Their system is based on marks like lines and bars. To create animations, the author defines the data for the initial state and the data for the final state. Their tool then interpolates between the two to generate an animated transition.

Very recently, Tableau, one of the largest providers of commercial visualization software has released a preview of their future implementation of animated transitions [95]. In their concept, animations are predetermined by the change in visualization parameters (sec. 2.1.1) and not configurable by the author.

The tool that most closely addresses the problems of narrative visualization is *Ellipsis* by Satyanarayan and Heer [96]. Ellipsis is a tool to create narrative visualization through a graphical interface. It wraps existing visualizations and adds an *annotation* layer, *parameters* for limited interaction (see section 1.1) and a layer for defining *transitions* between charts. Ellipsis thus effectively decouples the narrative structure from the individual charts and lets authors quickly explore alternative narrative structures.

11.2.3. Performance optimization

Heer et al. [93] mention another advantage of DSLs for visualization specification: they can easily be optimized without changing anything in the specification. The authors changed the implementation to use multiple threads and doubled their rendering speed without changing anything in the specification of the visualization.

11.3. Declarative syntax

Based on the requirements described in section 11.1 and the review of existing tools, we decided to design a domain-specific language (DSL) for narrative visualization with animated transitions. A DSL allows for *rapid prototyping* by making the definition of charts and transitions quick. Yet it can provide a lot of flexibility in terms of annotations and in how characters should morph between the charts. Finally, slow animations can be optimized without changing the visualizations that were already created for the experiment.

11.3.1. Anatomy of a chart

The first part of the DSL is concerned with defining the basic charts. The syntax is based on JSON and strongly inspired by Vega-Lite [94]. Vega-Lite is more powerful in inferring a multitude of chart types based solely on the configuration than our proposed solution. As we only required a very limited number of chart types (sec. 11.1), we resorted to the “Chart typology”-approach which is less flexible but much easier to implement. Two other differences of the proposed syntax to Vega-Lite are the focus on *characters* and the possibility to *annotate* them. An overview of the syntax is presented in figure 28.

```
"charts": [{
  "name": "D",
  "type": "stackedArea",
  "data": "../data/D.csv",
  "annotations": [{
      "name": "How the share of different energy sources
              has changed in California",
      "offset": {"left": 0}
    }
  ],
  "axes": [
    {
      "name": "y",
      "field": "share",
      "type": "quantitative",
      "domain": [0, 100],
      "annotations": [{
        "name": "Share of each energy source",
        "offset": {"top": -30, "left": 0}
      }]
    },
    {
      "name": "x",
      "field": "year",
      "type": "temporal",
      "domain": ["2002-01-01", "2017-01-01"],
      "ticks": 8
    }],
    "cast": {
      "field" : "source",
      "characters": []
    },
    "design":
      {
        "font":
          {
            "family": "Eczar",
            "size": 13
          },
        "margin": {"top": 160, "right": 100}}
}
```

source	year	share
solar	2002-01-01	0.302
crude oil	2002-01-01	1.067
coal	2002-01-01	1.266
wind	2002-01-01	2.069

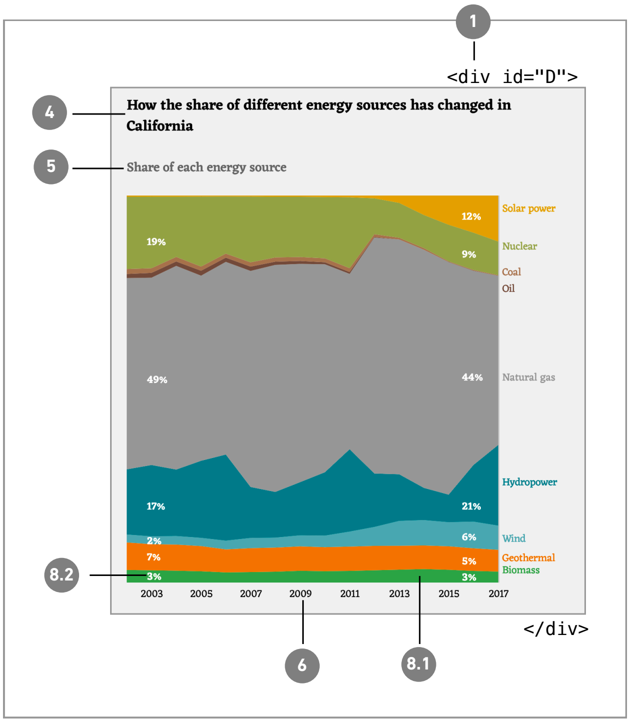


Figure 28: The code on the left leads to the generation of the chart on the right. On the top right an example of the type of data that is being used is shown.

- ① Each chart has a `name` which is used to identify it. If there is an HTML element with an ID that corresponds to this name it will be used to render the chart there. The chart dimensions as well es its position on the page are therefore completely defined by the layout of the surrounding page.
- ② The `type` defines the basic chart type that should be used. Behind the scenes, each chart type is implemented as a subclass of an abstract `Chart`-class that implements its own `draw`-method.
- ③ The path to a `CSV`-file with data. The syntax does not foresee any kind of data transformation functionality and expects the data to be in the right format. This is because the tool is presentation-oriented and there is no need to explore the data through filtering or other transformations.
- ④ The chart title is implemented as an `annotation`. Annotations can be bound to different elements throughout the DSL and can be positioned relative to their `start` or `end` through the `offset`-property.

- 5 In the axis-definitions, different attributes (`field`) of the data are mapped to different axes. The axes are identified by their `name` which can differ depending on the chart `type`. Slope charts, for example, have three axes: `x`, `from` and `to`. The advantage of defining each axis explicitly like this is, that, again, `annotations` can be bound to them which provided a lot of flexibility.
- 6 The `type`-property of an axis defines how the corresponding data should be parsed. The data domain can either be continuous numbers (`quantitative`), ordered discrete values (`ordinal`), unordered discrete values (`categorical`), or timestamps (`temporal`) [5, p. 21], [94]. The `domain`-property defines the corresponding start and end values of the axis. Often, the minimal and maximal values in the data are used to determine the start and end of an axis. But when using visualization for storytelling, sometimes the author wants values to “overshoot” the axis, or to fix the axis at a certain domain for dramatic or clarity reasons, which is why we allowed explicit control of this parameter.
- 7 Coherent with the model developed in section (>>Perception-oriented classification), we look at narrative visualization through the lens of characters. This is why our DSL contains an explicit declaration of the whole `cast` of characters in each chart. We assume that individual characters are identified by a `categorical` attribute in the data which is defined in the `field`-property.
- 8 Because characters have such high importance, they are defined individually (figure 29):

```
"characters": [
  {
    "name": "biomass",       8.1
    "color": "#53A353",
    "annotations": [
      {
        "name": "Biomass",
        "offset": {"top": -8}
      },
      {
        "name": "3%",
        "anchor": "start",    8.2
        "offset": {"left": 20, "top": 2},
        "class": "percent"
      },
      {
        "name": "3%",
        "anchor": "end",
        "offset": {"left": -50, "top": 1},
        "class": "percent"
      }
    ]
  },
{ ... }  8.3
]
```

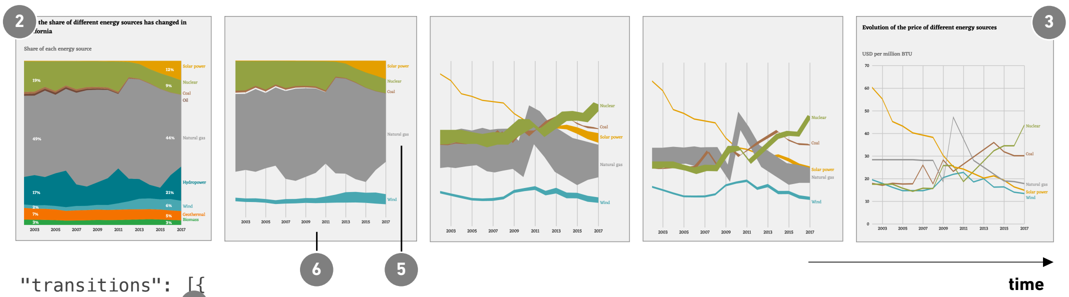
Figure 29: An example of how individual characters are defined.

- 8.1 In the three chart types that were implemented, characters are uniformly distinguished by `color` which is also what is often used in the narrative visualizations that were analyzed in section sec. 4.3. Other options like symbols or textures are naturally imaginable. The necessary data to `draw` each character is found via its `name` property.

- 8.2 Shows an example of the use of multiple `annotations`. Two are used to indicate the initial and the final share of each energy source, the third one to label the energy source itself. This approach has proven to provide a lot of flexibility. Especially when coupled with CSS `classes` that make individual styling of characters possible.
- 8.3 Such flexibility comes at the price of verbosity, as each character needs to be specified individually.
- 9 Finally, some basic visual properties of the chart can be defined in the `design` section of the specification.

11.3.2. Anatomy of a transition

Together with the `director` (sec. 11.3.3), this is the main contribution in terms of software architecture. Because we have identified character and attribute changes as the main concern for transitions, the DSL focuses on them. Characters as well as axes (which represent attributes), can be mapped between two charts which will create a third transition chart that interpolates between them (see figure 30).




```
"transitions": [{
  "name": "DE", 1
  "from": "D",  2
  "to": "E",    3
  "type": "morphing", 4
  "characters": [
    {"from": "solar", "to": "solar"}, 5
    {"from": "natural gas", "to": "natural gas"},
    {"from": "coal", "to": "coal"},
    {"from": "nuclear", "to": "nuclear"},
    {"from": "wind", "to": "wind"}
  ],
  "axes": [
    {"from": "x", "to": "x"} 6
  ]
}]
```

Figure 30: An example of how a *morphing* animation between individual characters from two charts can be declared.

- 1 A transition is not treated very differently from a chart. It also has a `name` and is rendered into an HTML element with the corresponding ID.
- 2 + 3 A transition chart is solely defined by the charts it transitions between. This is done through the `from` and `to` properties.
- 4 A `morphing` transition is used when characters are shared between charts (sec. 4.3). In other cases, the animation is a simple `fade`.
- 5 While some transition types like "Explore attributes" have a one-to-one relationship between the characters of the two charts, transition types like "Split characters" and "Merge characters" have a one-to-many or many-to-one relationship (sec. 4.3). This is modeled by mapping the same character in the `from` property to multiple characters in the `to` property or the inverse.
- 6 The system also supports axis interpolation (see design discussion in appendix J). Any axis in the initial chart can be mapped to any axis in the final chart. This particularly makes "Reconfigure" transitions possible.

11.3.3. Directing all of them

Finally, the `director` contains a kind of "scroll-timeline" of charts. Based on this it will show, hide and interpolate the appropriate charts at certain scroll positions. An example can be found in figure 31.

```
"director": {
  "name": "ESAD",  1
  "steps": [
    {"from": -1000, "to":100, "draw":"D"},
    {"from": 100, "to":900, "draw":"DE"},  2
    {"from": 900, "to":10000, "draw":"E"}
  ]
}
```

Figure 31: An example of how the sequence of charts and transitions can be defined in the `director`.

- 1 The director's `name` is used in the experiment to identify the mini-story and the configuration the participant was looking at.
- 2 The steps define a range of scroll positions and a chart or transition that should be displayed within this range. The `-1000` value is in there because some browsers permit scrolling above the start of the page, which would hide the first chart.

11.4. Program generation and rendering process

1. In a first step, the system generates objects for all the charts with their axes, characters, and annotations from the specification (sec. 11.3). Wherever something is not defined, it assumes a sensible default if possible or explains the failure otherwise.

2. After the charts, the system generates the objects who represent the transitions which reference the chart objects.
3. Finally, a director object is generated with references to all the charts and transitions. The director continually checks if the viewer has been scrolling.
4. When the scroll position changes, the director calls the `draw`-method on the charts/transitions that should be visible at this position. It calls the `hide`-method on all the others. If at the current position, a transition should be displayed, the `draw`-method is called with an additional `position` parameter that indicates the state of the interpolation.
5. This triggers a rendering cascade where every object also calls the `draw` methods of its children to render the final display. This corresponds to the *render*-pattern described by Heer et al. [97].
6. Objects of transition charts [characters, axes, annotations] implement their own interpolation methods and use them for rendering interpolated states.
7. The `draw`-methods generate the necessary SVG-code and append it to an SVG that will finally be displayed in the HTML element for the chart.

Our implementation of the described program generator and rendering process is referenced in appendix I

11.5. Interpolation

This section contains a few remarks on how the interpolation described above is being handled in our implementation. While previous work has usually interpolated between visualization parameters [93], [96], our system interpolates between SVG-shapes. Thanks to this, our system can generate animated transitions between two completely different chart types. The only condition is that a character needs to be represented by a single, closed SVG shape. Other systems only permit transitions “within” the same chart.

One problem when interpolating directly between SVG-shapes is that they need to have a one-to-one correspondence between their anchor points. If this is not the case, there needs to be some method to add points to the simpler shape. For this, we have used an algorithm that duplicates points of the simpler shape that are closest to the matching points in the complex shape [98]. This has produced visually good results in all our cases. But it had the downside of making the calculation of interpolated states slower.

11.6. Performance

As presented in the results, the drawing performance was not above the defined threshold for 7 of the participants (sec. 8). The distribution of the drawing performance over all the sessions with animated transitions in the experiment can be seen in figure 32.

Animation on the web typically has very mediocre performance. But there are a few strategies to make it faster which will be discussed in this section.

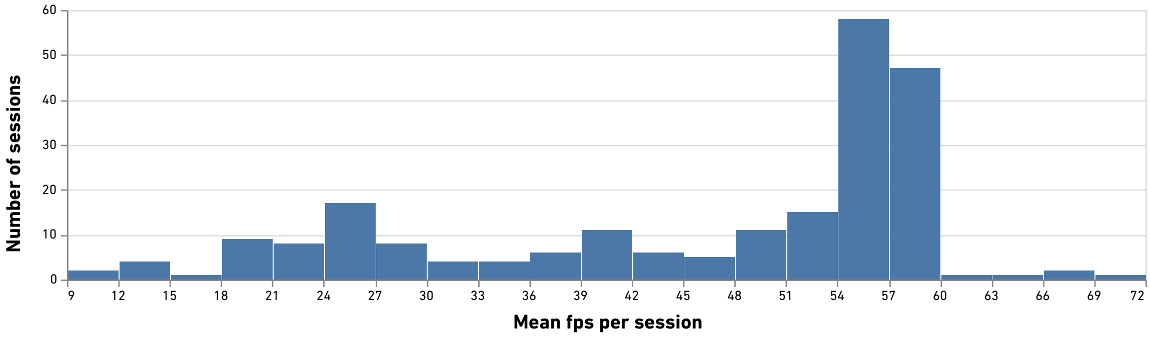


Figure 32: The mean frame rates by session in the experiment.

Synchronize repaint timing There is no connection between the time when a “scroll” event is being fired on the DOM and when the browser is ready to repaint the page. But thanks to the `requestAnimationFrame`-method, we can register a callback function that the browser will execute before the next repaint. In this callback, the director checks if the vertical position of the viewport has changed since the last call. If so, the DOM for this position is recalculated. Thanks to this, such a recalculation only ever happens when it will actually be painted by the browser. Typically, this is at 60fps but the rate is automatically toggled when calculations and repainting take more time. This approach has been implemented

Conditional redrawing Every component is responsible for drawing itself. `draw` methods could therefore easily be modified to only manipulate the DOM when something has changed (for example based on a “dirty”-flag). Based on our performance tests, we strongly suspect that browsers already implement this mechanism themselves and do not repaint elements whose subtree has simply been replaced with an identical subtree. We, therefore, did not implement this approach.

Parameter-based caching Because the system is highly decoupled, expensive calculations could easily be cached based on the input parameters. One very good candidate for this is the calculation of additional points for SVG shapes that will be interpolated. Because in our tests, the frame rates have always been much better than our baseline of 20fps, we did not implement this optimization. This decision was unfortunate as quite a few participants suffered from subpar performance (see figure 32 and section 8)

Canvas or WebGL Finally, we could handle the rendering of the pixels directly by using HTML `canvas` instead of SVG. But with current browsers and on current hardware, rendering performance has been shown to be about equal [99]. WebGL, on the other, hand is faster but would have made the implementation considerably more complicated [100].

11.7. Logging

Every interaction with the experiment was logged (sec. 7.4.3). Whenever the director detected scrolling, a log entry was created. Every keystroke too was recorded. And naturally, the participants answers where recorded. Logging is also the director's responsibility. Log entries are appended to a list which is sent to a server every 5 seconds as a CSV to minimize the potential for data loss. The data is also sent when the participant submits the answer form.

Every entry stores a timestamp, an identifier for the participant and the current mini-story, the URL, the scroll position and information about the participant's browser like window size or user agent. The full documentation can be found in appendix G.

11.8. Discussion

The solution presented in chapter 11.3 might look simple, even obvious to the reader. But our first attempts to implement animated transitions resulted in a tangled mess of spaghetti code and attest that the problem does not have a straightforward solution. We conclude that the simplicity of the proposed solution is, therefore, more a sign of conceptual clarity than a sign for the mundanity of the problem.

Part V.
Conclusion

12. Conclusions and future work

In the present work, we have asked if animated transitions help readers understand transitions in narrative visualization.

We have found that none of the existing literature proposes a concept of how readers might readers interpret transitions in narrative visualization. We have addressed this gap in chapter 4 by reconceptualizing three existing views with a focus on narrative visualization. We have successfully applied this reconceptualization to examples from practice which has resulted in the identification of 9 transition types. These types map well to classifications that have been proposed by other authors (see section 2.1) and clarify them. A future study should explore this formulation more in-depth. First, it should further research how well the *character*-metaphor represents how readers truly interpret narrative visualization. This might also lead to additional hints on *congruent* animation. Second, it should examine if and how the amount of change in *characters*, *attributes* and *context* is related to the difficulty of a transition. This might lead to a more formal definition of the *transition cost* (sec. 2.3). Such a definition would be highly useful for designers. They could use it to keep the *transition cost* within certain boundaries and ensure that a majority of readers would understand the stories presented.

In our review of the literature on animation (sec. 3) we have discussed the *congruence* and the *apprehension*-principle. The existing research indicates that *congruence* is much more important in practice than *apprehension*. While techniques that aim to increase apprehension often fail, animations that are close to the movement of objects in the real world consistently improve understanding. A future study should clarify how these two principles are linked to different levels of human perception. This would support designers in deciding when animation is useful and when it should be avoided.

The results of our experiment indicate that animated transitions do not support readers in understanding transitions. This answers the main research question and confirms previous findings [17]. But we have found in many discussions that it is rather unintuitive for visualization designers. It is therefore clear that these findings might be criticized based on weaknesses in the research method. The discussion of the limitations (sec. 10) gives many indications on how a better experiment could be designed. Especially, by considering the different transition types and how they can benefit from different strengths of animation (sec. 5.2). We plan to implement and conduct such an experiment in the future.

We were also surprised to find how little evidence there was in visualization research for the claim that animation increases engagement. While our results support previous research [54], much more work is needed to get a clear picture. Especially, visualization researchers should include validated scales for engagement more often in their studies.

Finally, we have presented a syntax to define animated transitions declaratively. The strength of our approach is that it is directly linked to the conceptual understanding and the transition types presented in chapter 4. It is therefore directly applicable to the design of narrative visualization. If future research would find that animated transitions provide benefits for certain transition types, our work on declaratively defining transitions, therefore, has potential. Because it's syntax is inspired by the very popular Vega-Lite [94], it might be adapted to integrate into Vega-Lite as a plugin. This would make it easy to link two charts defined in Vega-Lite via an animated transition. This would result in a useful design tool for narrative visualization.

References

[1] D. Reinsel, J. Gantz, and J. Rydning, "Data Age 2025," IDC, 2017.

[2] S. Marsh, "Neurotechnology, Elon Musk and the goal of human enhancement," *The Guardian: Technology*.

[3] J. Bischof, "A surprising number of people trust AI to make better policy decisions than politicians," *Quartz*. [Online]. Available: https://qz.com/1576057/could-ai-make-better-policy-than-politicians/. [Accessed: 30-Aug-2019].

[4] E. R. Tufte, *The visual display of quantitative information*. Cheshire, CT: Graphics Press, 2001.

[5] T. Munzner and E. Maguire, *Visualization analysis and design*. Boca Raton, FL: CRC Press, 2015.

[6] J. Stasko, "Tweet." 22-Aug-2019.

[7] L. Gamio, "Urban and rural America are becoming increasingly polarized," *Washington Post*, 17-Nov-2016.

[8] E. Barry, "The Russia Left Behind," *The New York Times*, 13-Oct-2013.

[9] J. Dworkin and I. Blinderman, "Why the tech sector may not solve America's looming automation crisis," *The Pudding*, Aug-2018.

[10] L. Borgenheimer, P. Blickle, J. Stahnke, S. Venohr, and C. Bangel, "German Unification: A Nation Divided," *Die Zeit*, Hamburg, 19-Nov-2014.

[11] M. Stabe, "Why the FT creates so few clickable graphics," *Financial Times*, 03-Oct-2016.

[12] A. Tse, "Why We Are Doing Fewer Interactives," presented at the Malofiej Graphics, 01-Jan-2016.

[13] J. Boy, L. Eveillard, F. Detienne, and J.-D. Fekete, "Suggested Interactivity: Seeking Perceived Affordances for Information Visualization," *IEEE Trans. Visual. Comput. Graphics*, vol. 22, no. 1, pp. 639–648, Aug. 2015.

[14] N. H. Riche, C. Hurter, N. Diakopoulos, and S. Carpendale, Eds., *Data-driven storytelling*. Boca Raton, Florida: CRC Press/Taylor & Francis Group, 2018.

[15] P. W. Thorndyke, "Cognitive structures in comprehension and memory of narrative discourse," *Cognitive Psychology*, vol. 9, no. 1, pp. 77–110, Jan. 1977.

[16] B. Tversky, J. B. Morrison, and M. Bétrancourt, "Animation - can it facilitate?" *Int. J. Hum.-Comput. Stud.*, vol. 57, no. 4, pp. 247–262, Jan. 2002.

[17] S. Berney and M. Bétrancourt, "Does animation enhance learning? A meta-analysis," *Computers & Education*, vol. 101, pp. 150–167, Oct. 2016.

[18] E. Segel and J. Heer, "Narrative Visualization - Telling Stories with Data." *IEEE Trans. Vis. Comput. Graph.*, Jan. 2010.

[19] J. Heer and G. G. Robertson, "Animated transitions in statistical data graphics," *IEEE Trans. Visual. Comput. Graphics*, vol. 13, no. 6, pp. 1240–1247, Nov. 2007.

[20] J. S. Yi, Y. ah Kang, J. T. Stasko, and J. A. Jacko, "Toward a Deeper Understanding of the Role of Interaction in Information Visualization." *IEEE Trans. Vis. Comput. Graph.*, vol. 13, no. 6, pp. 1224–1231, Jan. 2007.

[21] D. Fisher, "Animation for Visualization: Opportunities and Drawbacks," in *Beautiful Visualization*, 2011, p. 24.

[22] anbync, "The Space Between." [Online]. Available: https://observablehq.com/@anbnyc/the-space-between. [Accessed: 29-Jul-2019].

[23] T. J. Jankun-Kelly, K.-L. Ma, and M. Gertz, "A Model and Framework for Visualization Exploration." *IEEE Trans. Vis. Comput. Graph.*, vol. 13, no. 2, pp. 357–369, Jan. 2007.

[24] Google, "Google Maps." [Online]. Available: https://www.google.com/maps. [Accessed: 05-Aug-2019].

[25] S. McCloud, *Understanding comics*. 1993.

[26] P. Baudisch, D. Tan, and M. Collomb, "Phosphor: Explaining Transitions in the User Interface," *Citeseer*, Jan. 2006.

[27] F. Amini, N. H. Riche, and B. Lee, "Understanding data videos: Looking at narrative visualization through the cinematography lens," *dl.acm.org*, Jan. 2015.

[28] D. Y. M. Badawood, "Narrative Construction in Information Visualization," p. 220, 2015.

[29] B. Bach, N. Kerracher, K. W. Hall, S. Carpendale, J. Kennedy, and N. Henry Riche, "Telling Stories about Dynamic Networks with Graph Comics," in *CHI '16*, 2016, pp. 3670–3682.

[30] N. Cohn, "The limits of time and transitions: Challenges to theories of sequential image comprehension," 2010.

[31] Wikipedia, "Five Ws," *Wikipedia*. 23-Jul-2019.

[32] J. Hullman, S. M. Drucker, N. H. Riche, B. Lee, D. Fisher, and E. Adar, "A Deeper Understanding of Sequence in Narrative Visualization." *IEEE Trans. Visual. Comput. Graphics*, vol. 19, no. 12, pp. 2406–2415, Jan. 2013.

[33] E. Roston and B. Migliozzi, "What's Really Warming the World? Climate deniers blame natural factors; NASA data proves otherwise," *Bloomberg*, 24-Jun-2015.

[34] M. Lambrechts, "Why EU Regions are Redrawing Their Borders," *The Pudding*, Apr-2019.

[35] L. Nathaniel, "How five of the worst schools in Florida wound up in Pinellas County," *Tampa Bay times*, 12-Aug-2015.

[36] G. Aisch, R. Gebeloff, and K. Quealy, "Where We Came From and Where We Went, State by State," *The New York Times: The Upshot*, 14-Aug-2014.

[37] D. Y. Fernholz Tim, "Interactive graphic: See every satellite orbiting earth," *Quartz*, 21-Dec-2015. [Online]. Available: https://qz.com/296941/interactive-graphic-every-active-satellite-orbiting-earth/. [Accessed: 28-Aug-2019].

[38] M. Brehmer and T. Munzner, "A Multi-Level Typology of Abstract Visualization Tasks." *IEEE Trans. Vis. Comput. Graph.*, vol. 19, no. 12, pp. 2376–2385, Jan. 2013.

[39] M. Gleicher, "Considerations for Visualizing Comparison," *IEEE Trans. Visual. Comput. Graphics*, vol. 24, no. 1, pp. 413–423, 2018.

[40] D. A. Szafir, S. Haroz, M. Gleicher, and S. Franconeri, "Four types of ensemble coding in data visualizations," *Journal of Vision*, vol. 16, no. 5, pp. 11–11, Mar. 2016.

[41] A. Wong, W. Leahy, N. Marcus, and J. Sweller, "Cognitive load theory, the transient information effect and e-learning," *Learning and Instruction*, vol. 22, no. 6, pp. 449–457, Dec. 2012.

[42] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information." *Psychological Review*, vol. 63, no. 2, 1956.

[43] G. A. Alvarez and P. Cavanagh, "The Capacity of Visual Short-Term Memory is Set Both by Visual Information Load and by Number of Objects," *Psychol Sci*, vol. 15, no. 2, pp. 106–111, Jan. 2004.

[44] F. Chevalier, N. H. Riche, C. Plaisant, A. Chalbi, and C. Hurter, "Animations 25 Years Later," in *AVI '16*, 2016, pp. 280–287.

[45] F. Chevalier, P. Dragicevic, and S. Franconeri, "The Not-so-Staggering Effect of Staggered Animated Transitions on Visual Tracking." *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 12, pp. 2241–2250, Jan. 2014.

[46] R. A. Rensink, "Change Detection," *Annual Review of Psychology*, vol. 53, no. 1, pp. 245–277, 2002.

[47] A. Chalbi, "Understanding and designing animations in the user interfaces," 2018.

[48] E. Wu, L. Jiang, L. Xu, and A. Nandi, "Graphical Perception in Animated Bar Charts," Mar. 2016.

[49] B. B. Bederson and A. Boltman, "Does Animation Help Users Build Mental Maps of Spatial Information?" p. 10.

[50] C. Ware and R. Bobrow, "Motion to support rapid interactive queries on node–link diagrams," *ACM Trans. Appl. Percept.*, vol. 1, no. 1, pp. 3–18, Jul. 2004.

[51] M. Bétrancourt, P. Dillenbourg, and L. Clavien, "Display of Key Pictures from Animation: Effects on Learning," in *Understanding Multimedia Documents*, Chapter 4 vols., Boston, MA: Springer, Boston, MA, 2008, pp. 61–78.

[52] N. Kadaba, N. Kadaba, N. Kadaba, N. Kadaba, P. Irani, and J. Leboe, "Visualizing Causal Semantics Using Animations," *IEEE Trans. Visual. Comput. Graphics*, vol. 13, no. 6, pp. 1254–1261, Nov. 2007.

[53] Y. Kim, M. Correll, and J. Heer, "Designing Animated Transitions to Convey Aggregate Operations," *Computer Graphics Forum*, vol. 38, no. 3, pp. 541–551, Jun. 2019.

[54] F. Amini, N. H. Riche, B. Lee, J. Leboe-McGowan, and P. Irani, "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*, 2018, pp. 1–9.

[55] M. Hassenzahl and A. Monk, "The Inference of Perceived Usability From Beauty," *Human-Comp. Interaction*, vol. 25, no. 3, pp. 235–260, Jul. 2010.

[56] B. Bach, E. Pietriga, and J.-D. Fekete, "GraphDiaries: Animated Transitions andTemporal Navigation for Dynamic Networks," *IEEE Trans. Visual. Comput. Graphics*, vol. 20, no. 5, pp. 740–754, Jan. 2014.

[57] P. I. M. Shanmugasundaram, "Can Smooth View Transitions Facilitate Perceptual Constancy in Node-Link Diagrams?" pp. 1–8, Mar. 2007.

[58] M. Shanmugasundaram and P. Irani, "The effect of animated transitions in zooming interfaces." *AVI*, p. 396, Jan. 2008.

[59] S. L. Franconeri, S. V. Jonathan, and J. M. Scimeca, "Tracking Multiple Objects Is Limited Only by Object Spacing, Not by Speed, Time, or Capacity," *Psychol Sci*, vol. 21, no. 7, pp. 920–925, Jun. 2010.

[60] S. Yantis, “Multielement visual tracking: Attention and perceptual organization,” *Cognitive Psychology*, vol. 24, no. 3, pp. 295–340, Jul. 1992.

[61] B. J. Scholl and Z. W. Pylyshyn, “Tracking Multiple Items Through Occlusion: Clues to Visual Objecthood,” *Cognitive Psychology*, vol. 38, no. 2, pp. 259–290, Jan. 1999.

[62] M. Chan and J. Black, *When can animation improve learning? Some implications on human computer interaction and learning.* Association for the Advancement of Computing in Education (AACE), 2005.

[63] B. S. Hasler, B. Kersten, and J. Sweller, “Learner control, cognitive load and instructional animation,” *Appl. Cognit. Psychol.*, vol. 21, no. 6, pp. 713–729, Jan. 2007.

[64] J. Lasseter, “Principles of traditional animation applied to 3D computer animation.” *SIGGRAPH*, pp. 35–44, Jan. 1987.

[65] P. Dragicevic, A. Bezerianos, W. J. P. of the, and 2011, “Temporal distortion for animated transitions,” *dl.acm.org*, Jan. 2012.

[66] S. McKenna, N. Henry Riche, B. Lee, J. Boy, and M. Meyer, “Visual Narrative Flow: Exploring Factors Shaping Data Visualization Story Reading Experiences,” *Computer Graphics Forum*, vol. 36, no. 3, pp. 377–387, Jun. 2017.

[67] E. Badger, C. C. Miller, A. Pearce, and K. Quealy, “Extensive Data Shows Punishing Reach of Racism for Black Boys,” *The New York Times: The Upshot*, 19-Mar-2018.

[68] D. Bauer, “Here’s why men on earth outnumber women by 66 million,” *Quartz*, 18-Oct-2016.

[69] K. Collins, A. Pearce, and D. Armstrong, “Why Measles May Just Be Getting Started,” *Bloomberg*, 06-Feb-2015.

[70] B. Victor, “Up and Down the Ladder of Abstraction,” Nov-2011. [Online]. Available: http://worrydream.com/LadderOfAbstraction/. [Accessed: 09-Aug-2019].

[71] M. Klein, “How Americans Die,” *Bloomberg*, Apr. 2014.

[72] M. Daniels, “Using Historical NBA Data To Rank the Most Unlikely Comebacks,” *The Pudding*. [Online]. Available: https://pudding.cool/2017/03/comeback/index.html. [Accessed: 30-Aug-2019].

[73] J. Hernandez and Q. Bui, “The American Dream Is Alive. In China.” *The New York Times: World*, 18-Nov-2018.

[74] B. Casselman, “Strong Hiring Still Isn’t Bringing Pay Raises,” *FiveThirtyEight*, 06-Mar-2015. [Online]. Available: https://fivethirtyeight.com/features/strong-hiring-still-isnt-bringing-pay-raises/. [Accessed: 30-Aug-2019].

[75] E. Klein, “Obama on American politics and economy: The Vox conversation,” *Vox*, 2014. [Online]. Available: https://www.vox.com/a/barack-obama-interview-vox-conversation/obama-domestic-

policy-transcript. [Accessed: 30-Aug-2019].

[76] S. Franconeri, “Thinking with Data Visualizations, Fast and Slow,” presented at the OpenVis Conference 2018, Paris, 01-Aug-2018.

[77] none, “Visual Thinking Lab.” [Online]. Available: http://visualthinking.psych.northwestern.edu/. [Accessed: 30-Aug-2019].

[78] None, “Amazon Mechanical Turk.” [Online]. Available: https://www.mturk.com/. [Accessed: 20-Aug-2019].

[79] N. Popovich, “How Does Your State Make Electricity?” *The New York Times: Climate*, 24-Dec-2018.

[80] M. Mahajan, “Plunging Prices Mean Building New Renewable Energy Is Cheaper Than Running Existing Coal,” *Forbes*, 03-Dec-2018.

[81] H. L. O’Brien, P. Cairns, and M. Hall, “A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form,” *International Journal of Human-Computer Studies*, vol. 112, pp. 28–39, Apr. 2018.

[82] S. Ishihara, “Ishihara’s Test for Colour Deficiency: 38 Plates Edition,” 1917. [Online]. Available: https://www.color-blindness.com/ishiharas-test-for-colour-deficiency-38-plates-edition/. [Accessed: 14-Aug-2019].

[83] None, “Color blindness,” *Wikipedia*. 07-Aug-2019.

[84] J. Kaminski, “The science behind the service,” 06-Jun-2019. [Online]. Available: https://cloud.ibm.com/docs/services/tone-analyzer?topic=tone-analyzer-ssbts.

[85] M. Bostock, V. Ogievetsky, and J. Heer, “D^3 Data-Driven Documents,” *IEEE Trans. Visual. Comput. Graphics*, vol. 17, no. 12, pp. 2301–2309, Jan. 2011.

[86] J. Boy, R. A. Rensink, E. Bertini, and J.-D. Fekete, “A Principled Way of Assessing Visualization Literacy,” *IEEE Trans. Visual. Comput. Graphics*, vol. 20, no. 12, pp. 1963–1972, Jan. 2014.

[87] H. Mei, Y. Ma, Y. Wei, and W. Chen, “The design space of construction tools for information visualization: A survey,” *Journal of Visual Languages & Computing*, vol. 44, pp. 120–132, Feb. 2018.

[88] Y. Holtz and C. Healy, “From data to Viz.” [Online]. Available: data-to-viz.com. [Accessed: 15-Aug-2019].

[89] S. Ribecca, “The Data Visualisation Catalogue.” [Online]. Available: https://datavizcatalogue.com/. [Accessed: 19-Aug-2019].

[90] J. Schwabish and S. Ribecca, “Visual Vocabulary.” [Online]. Available: http://ft-interactive.github.io/visual-vocabulary/. [Accessed: 19-Aug-2019].

[91] M. Russo, A. Ferrari, and D. Perilli, “Microsoft Power BI – Visuals available,” Sep-2018. [Online]. Available: https://www.sqlbi.com/ref/power-bi-visuals-reference/. [Accessed: 19-Aug-2019].

[92] H. Wickham, *Ggplot2: Elegant graphics for data analysis*, Second edition. Cham: Springer, 2016.

[93] J. Heer and M. Bostock, “Declarative Language Design for Interactive Visualization,” *IEEE Trans. Visual. Comput. Graphics*, vol. 16, no. 6, pp. 1149–1156, Nov. 2010.

[94] A. Satyanarayan and D. Moritz, “Vega-lite: A grammar of interactive graphics,” *ieeexplore.ieee.org*, 2016.

[95] P. Isaacs and A. Luthy, “Tableau in motion – Tableau’s new native animations,” presented at the Tableau Conferences, 25-Oct-2018.

[96] A. Satyanarayan and J. Heer, “Authoring Narrative Visualizations with Ellipsis,” *Computer Graphics Forum*, vol. 33, no. 3, pp. 361–370, Jul. 2014.

[97] J. Heer and M. Agrawala, “Software Design Patterns for Information Visualization,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 853–860, Sep. 2006.

[98] P. Beshai, “Improving D3 Path Animation,” 14-Sep-2016. [Online]. Available: https://bocoup.com/blog/improving-d3-path-animation. [Accessed: 20-Aug-2019].

[99] T. Horak, U. Kister, and R. Dachselt, “Comparing Rendering Performance of Common Web Technologies for Large Graphs,” p. 2, 2018.

[100] D. E. Kee, L. Salowitz, and R. Chang, “Comparing Interactive Web-Based Visualization Rendering Techniques,” p. 2, 2012.

Part VI.
Appendices

A. Honesty declaration

Hereby, I declare that I have composed the presented work Lost in transition independently on my own and without any other resources than the ones indicated. All thoughts taken directly or indirectly from external sources are properly denoted as such.

Windisch,
August 30, 2019 Jonas Oesch

B. Corpus of narrative visualizations

The collection of 147 narrative visualizations from three collections. The collection is available on GitHub https://github.com/jonasoesch/lostintransition in the folder `supplementary/B-corpus`.

C. Analysis of narrative visualizations

The anlysis of of a selection of narrative visualizations to identify different transition types. The analysis is available on GitHub https://github.com/jonasoesch/lostintransition in the folder `supplementary/C-analysis`.

D. Stories

The mortality and energy story that were used in the experiment. The stories are linked on GitHub https://github.com/jonasoesch/lostintransition in the folder `supplementary/D-stories`.

E. Experiment questionnaire

The questionnaire that was presented after each mini-story. The questinnaire is available on GitHub https://github.com/jonasoesch/lostintransition in the folder `supplementary/E-questionnaire`.

F. Experiment survey

The survey that was presented to the participants at the end of the experiment. The survey is available on GitHub https://github.com/jonasoesch/lostintransition in the folder `supplementary/F-survey`.

G. Passive data collected in the experiment

The schema for the data that was collected passively during the experiment. The schema is available on GitHub https://github.com/jonasoesch/lostintransition in the folder `supplementary/G-passive-data`.

H. Data analysis

The jupyter notebooks that were used for the data analysis. The analysis is available on Github https://github.com/jonasoesch/lostintransition in the folder `supplementary/H-analysis`.

I. Implementation

An implementation of a program generator that interprets the described syntax and produces a narrative visualization with animated transitions. The implementation is linked on Github https://github.com/jonasoesch/lostintransition in the folder `supplementary/I-implementation`.

J. Design choices

Story selection

Selecting or designing the stories was one of the most delicate parts of the experiment design. The characters and the broader context needed to be familiar to a wide audience because visualization interpretation is highly dependent on context knowledge (compare sections 1.1 and 2.2.1) and the experiment did not provide an introduction to the topic. A story on paratransit (a special means of transport for disabled people in the U.S.) for example, was initially planned to be included but was soon removed because the topic was not familiar to most people.

The stories also had to avoid highly controversial topics because we feared that implicit reader knowledge might interfere with the interpretation of the visualizations (see section 2.2.1). We have seen this happen in both stories but it was less prevalent than we feared.

Finally, the story needed to be told mainly through the visualizations. This turned out to be to most constraining factor. In most existing narrative visualization we have found, that the textual narration was essential to understand the story. The story on the evolution of mortality was chosen exactly because it contained very little text in its original version. The story about the energy sources was specifically created in a way that we hoped would be self-explanatory with very little text.

Minimal textual narrative

Textual narrative was excluded from the experiment because it is a huge confounder. In typical narrative visualization, the story is presented through textual or audio narrative. The visualizations mostly serve to reinforce the point. But, when presented like this, relating the participants answers to any differences in the visualizations would be very difficult. On the other hand, charts are impossible to interpret without at least some text. We, therefore, decided to include labels as well as a chart title. But we made sure that nothing was hinting at a relationship in any of these. Each chart is completely self-contained and provides interesting information even without the other chart.

Visualization literacy

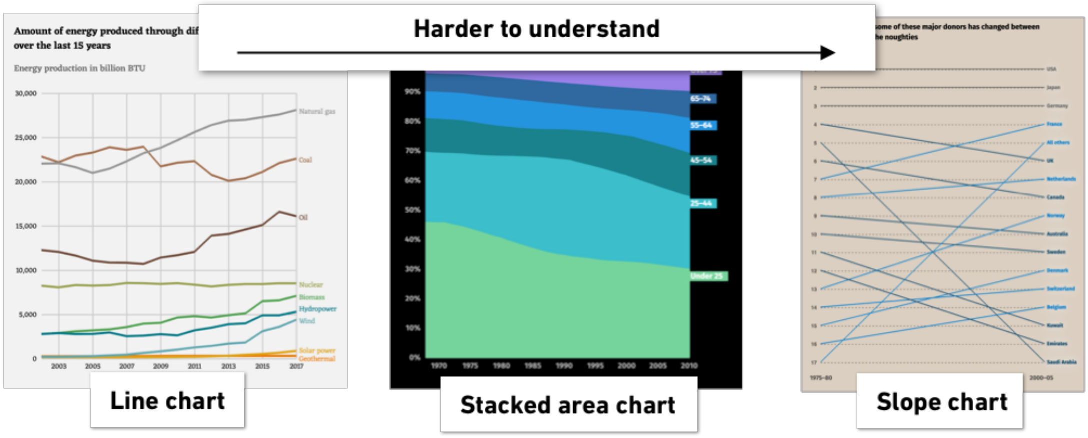


Figure 33: Different chart types and their difficulty according to our tests.

Previous studies have found that the majority of people are unable to correctly interpret complex visualizations. According to these authors, the "safe" visualizations are bar charts, line charts, scatterplots and maps. This finding is supported by our analysis of narrative visualization in section 4.3 who also almost exclusively use these simpler visualizations As we did not want participants to fail because they were unable to read the individual charts, we have decided to limit the experiment to line charts and stacked area charts. In our pilot studies, these chart types have been "safe", even though the stacked area chart posed problems to some of the participants in the experiment (see 8). One chart type that was excluded based on pilot data was the *slope chart*. (see 33)

Reader-controlled animation

Another factor to consider was the amount of reader control or interactivity. Prior work has demonstrated positive effects of animated transitions often involved higher levels of interactivity compared to the static transitions. But interactivity has been shown to have benefits in itself, for example for learning. All transitions were therefore designed to be totally reader controlled. By scrolling down, readers could advance to the next chart, by scrolling up, they could go back to the previous chart.

We also tried to avoid discussions about the proper duration of animated transitions by making the animations completely controllable through scrolling. Interestingly we have not found this to be common practice in our analysis of narrative visualizations. The most prevalent case is a fixed-duration animation that is triggered by scrolling to a certain point.

Animation design

Three different kinds of animated transitions were used throughout the experiment which were all concerned with supporting *object constancy* in different scenarios:

One-to-many A good example of this case is *Mortality D* (sec. 7.2.2). There, the animation needed to convey that the "causes of death" char in the second chart only concerned the "25–44" year old age group ctx . This is done through a *staged transition* (see section 3.2.3). We first highlight the "25–44" year old age group by hiding all the others and then splitting and morphing this character into the different causes of death. The same approach was applied in *Mortality B* (sec. 7.2.2), *Energy B* (sec. 7.2.3) and *Energy C* (sec. 7.2.3).

Many-to-one This case can only be found in *Mortality A* (sec. 7.2.2). Here the different age groups char are being morphed into a single line that represents "Everyone" char . After the morphing was finished, the characters ("Men" char and "Women" char) were shown.

Many-to-many This case is well illustrated by *Energy D* (sec. 7.2.3) where the marks for proportions attr of different energy sources char were morphed to represent the price evolution attr of these same energy sources char . Many-to-many animations are also being used in *Mortality C* (sec. 7.2.2) and *Energy A* (sec. 7.2.3).

Animated axis interpolation

In the first pilot study, we have included axis interpolation for some transitions. But they were excluded from the final experiment as they introduced another confounder and were not directly related to the research question. In our analysis of transitions in narrative visualization, we have found several different approaches to animating axis interpolation. A topic that would certainly merit further research.

Interpolation

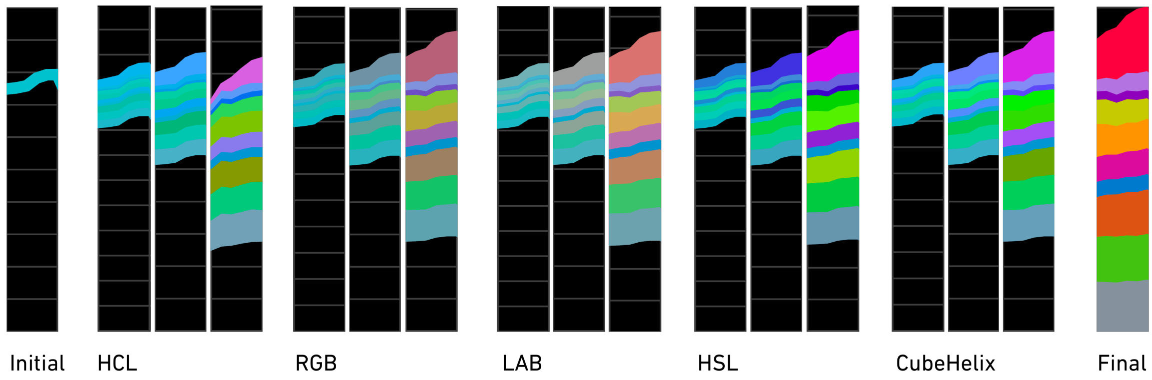


Figure 34: Different color interpolation mehthods compared.

For all interpolations, we implemented “slow-in-slow-out”-easing. For interpolating between colors, we used a perceptually uniform HCL-interpolation because it “intuitively looks right”. Notice in figure 34 how RGB and LAB tend to desaturate while HSL and CubeHelix tend to oversaturate, HCL strikes a good balance.

The End.