medium.com

Timelines are not always lines: An evaluation of different timeline shapes

Sara Di Bartolomeo

12-15 minutes

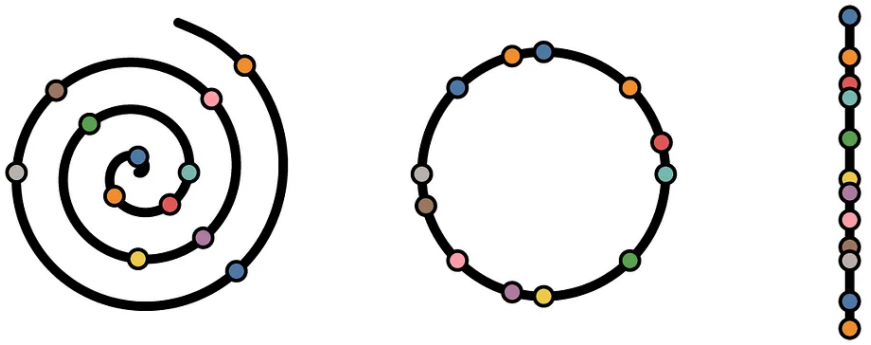
TL;DR: *When we create a timeline, we can use many different shapes to convey different meanings, or just to apply our creativity: it can be just a line, or it can be a circle, a spiral, or any other shape. How does the shape of a timeline affect how well people can read it? In recent research, we investigate how different timeline shapes, intended to convey the shape of the underlying data, affect how easily a user can interpret a timeline.*

Timelines come in many shapes and forms. When we think about a timeline, we usually think about a horizontal line:

> A linear timeline. Events are represented as colored dots on the line.

We know, though, that this is not the only way to represent a sequence of temporal events. We could use a line, a circle, a spiral,

or any other shape, and they would still convey how one event follows another:



Other possible options to represent temporal event data.

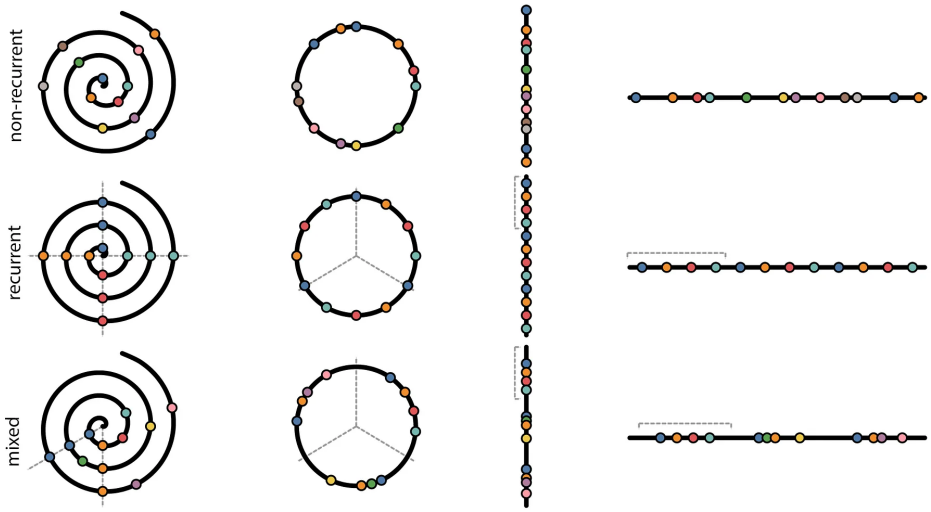
Let's consider the properties of the data that we want to represent. A dataset containing sequences of events can have events that repeat throughout the dataset, events that never repeat, or a mix of the two. It's easy to imagine real life datasets with these properties. We can define these categories:

- **non-recurrent** datasets do not contain repeating events. An example of this could be a historical timeline, in which each event has a date, and they don't repeat throughout history.
- **recurrent** datasets only contain repeating events. An example of this is the planting schedule for a farmer: each plant has a specific time frame in which it needs to be planted, and it reoccurs at about the same time every year.
- **mixed** datasets contain a mix of repeating and non repeating events. An example of this is the daily schedule of a person: it can have repeating events (e.g. classes, weekly recurring meetings) and non repeating events (e.g. parties, one time meetings).

We decided to use a made up historical timeline for a non-recurrent dataset, a made up planting schedule for a recurrent dataset, and a made up schedule of a person for a mixed dataset.

The question we asked in our research is: ***what timeline shape is best to represent each one of these types of temporal datasets?***

One could intuitively guess that the best timeline for a non-recurrent dataset could be a line, for a recurrent one could be a circle, while, perhaps, a spiral could do a good job representing a mixed dataset. In order to find out if this intuition was correct, we ran an experiment comparing all these different shapes combined with all these different data types.

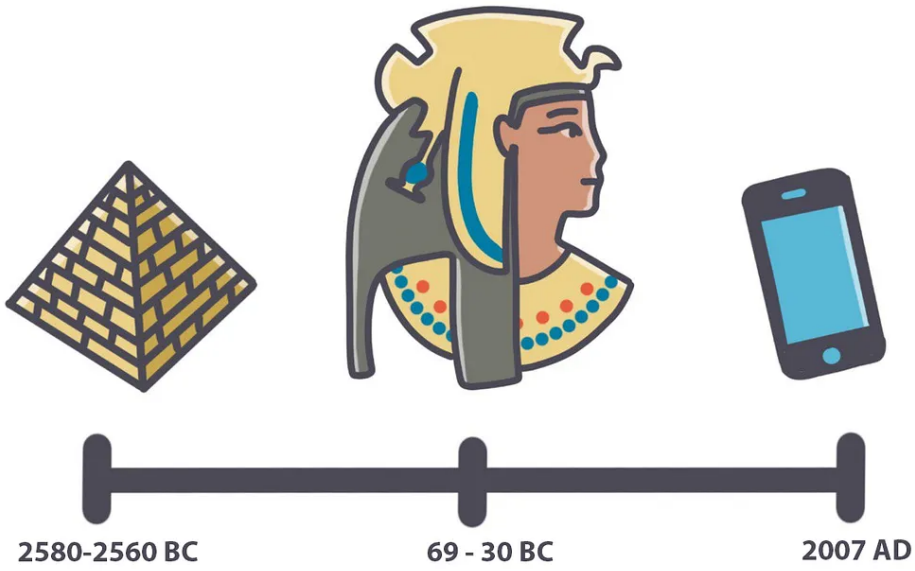


Which timeline is best for each dataset type? In this picture, you can see how events recur in mixed and recurrent datasets, where the intervals are shown with dashed lines. Dots of the same color represent the same type of event. In mixed datasets, some events recur, and some don't recur.

We first interviewed people who deal with timelines everyday — a history professor and a researcher working with temporal events in personal health informatics — in order to find out what people want to learn from timelines in different fields.

From our interviews, we selected four main tasks that are relevant for timelines:

- **when,** in which we want to figure out the date in which a specific event happened. One example of this would be: *When did the earthquake happen?*
- **what,** in which we want to figure out what event happened at a specific date. Example: *What happened in 1999?*
- **find,** in which we know both the event and the date, and want to find the location on the timeline. Example: *The earthquake happened in 1898. Find it on the timeline.*
- **compare,** in which we want to know the timing of an event relative to another one. Example: *Did Cleopatra live closer to the launch of the first iPhone or the construction of the Pyramids?*



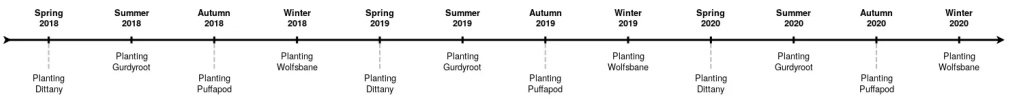
> Cleopatra lived about 2500 years after the construction of the pyramids, and 2037 years before the launch of the first iPhone. Therefore, surprisingly, she lived closer to the launch of the iPhone.

We then made up three datasets: a historical one for the non-recurrent category, a planting schedule for the recurrent category, and a person's schedule for the mixed one. In this way, we made sure the participants in our experiment would not have any previous knowledge about the data that they would be seeing.

The picture below shows our initial designs that we would use in our tests. Unfortunately, we soon found out that these colorful designs would introduce additional variables into the experiment, such as a color palette, or the orientation of the text, that could influence our results — while instead we just wanted to see the the effect of the shape. Therefore, we ended up ditching all of that and trying to achieve a design that would be as plain and simple and possible.

Our initial designs for the timelines were colorful and contained time spans. This gif showed how the same data could be displayed on many different shapes.

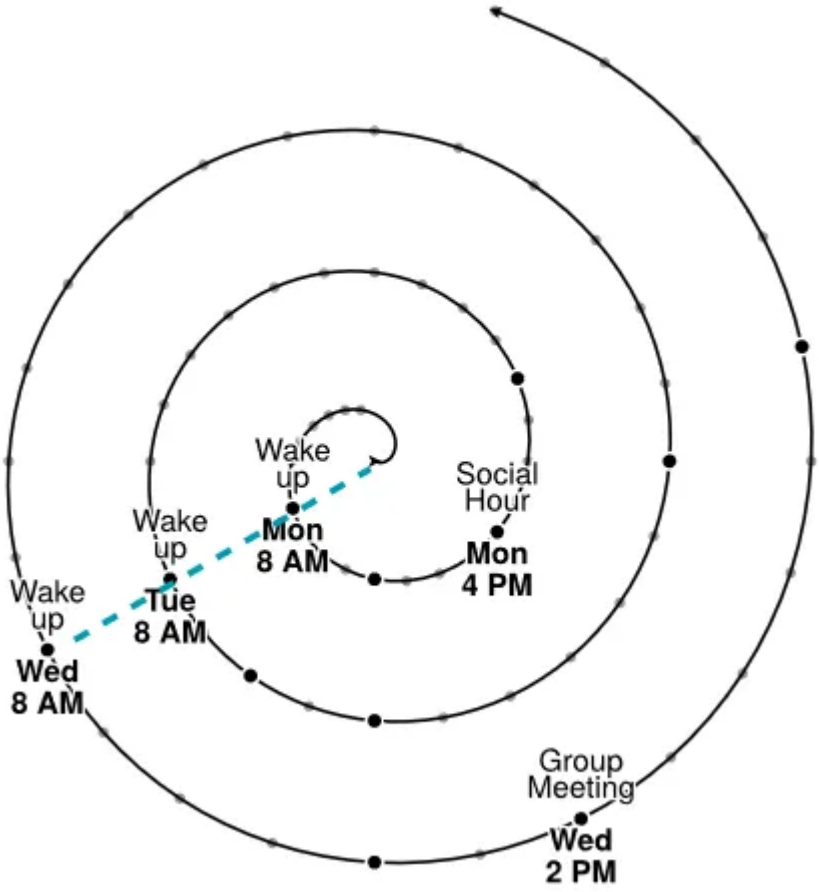
We also had to balance the number of data points to be shown in order to be clearly readable on many different screen sizes. We finally decided to settle for 12 data points on every timeline, all the text had to be horizontal, there would be no time spans, and the text would use a plain Arial font. Our horizontal timeline ended up looking like this:



Horizontal timeline with the planting dataset

The value of starting simple

The data and tasks we ended up using, as well as the designs, are extremely simple. We argue that to fully understand the effect of the shape of a timeline, we have to start from the simplest datasets and tasks.



> Here is an example of how a spiral timeline, with a mixed dataset (a schedule), could benefit the reader: the same hour can be at the same angle relative to the center for every round of the spiral, thus making it easier to find things that recur at the same time every day, while also making non-recurring events stand out. This is the design that we used in the experiment — the only difference being that in the actual experiment, the timelines contained 12 data points each.

Another reason behind the simplicity is that timelines are extremely widespread. Since almost everybody comes in contact with timelines on a daily basis, simple static visualizations play a valued role in communication.

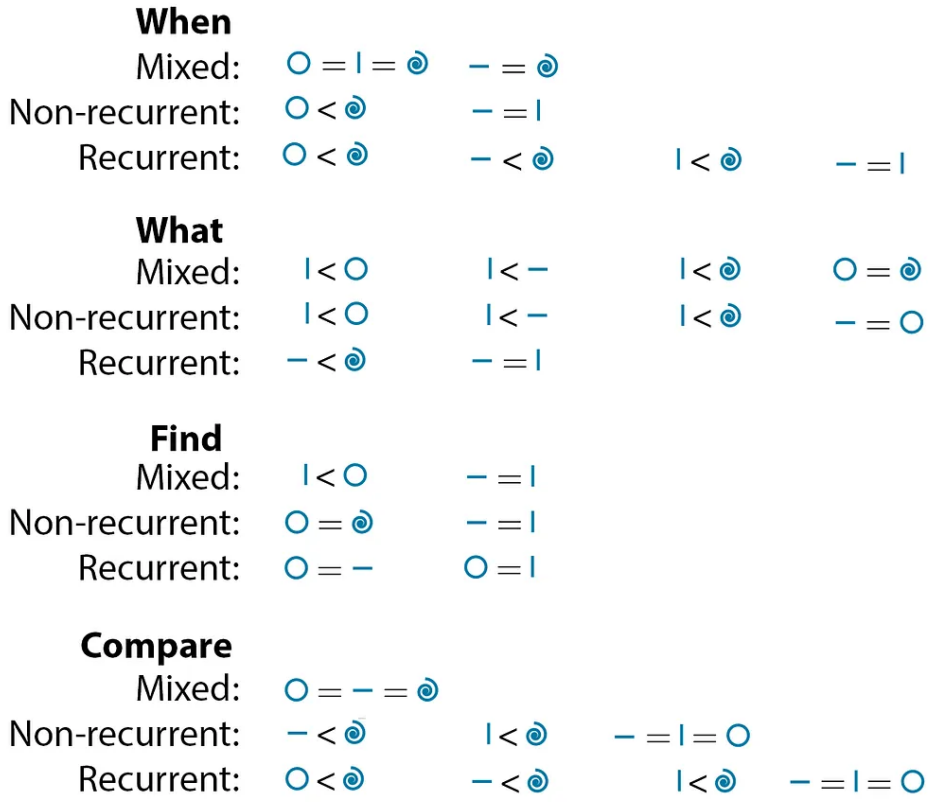
As our target population was vast and very diverse, we decided that using Mechanical Turk to recruit participants was our best option, as it would allow us to reach a very large audience. We gave each

participant a set of 12 questions, every time changing the dataset and the shape of the timeline we were showing them, so that they would observe each dataset type with each one of the timeline shapes.

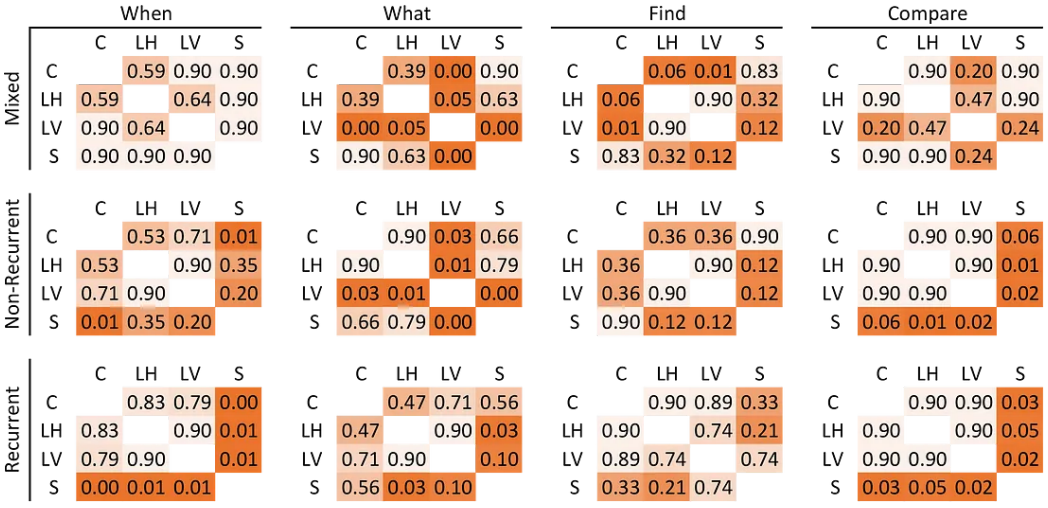
What we ended up finding out is that there is no overall winner, and no timeline is better than the others in every category. Some timelines work better than others in specific cases, while in some other cases, the choice of shape doesn't make any difference.

For example, in terms of **timing**, in the **what** task — where participants were asked to find out the date in which a specific event happened — vertical linear timelines have performed better than all the other shapes for the mixed and the non-recurrent dataset. In the **find** task, instead, where participants were asked to find an event knowing both the date and the name of the event, the choice is mostly unimpactful.

The results are a bit more complicated than this, so we made a picture summarizing our findings in terms of timing:



Here is a summary of our findings about timing. A “<” sign means that the timing on the first shape is, overall, less than the second one, while a “=” sign means that there is no timing difference between the two shapes.



Here is a chart of all of *the p*-values for the timing resulting from the experiment.

More in-depth graphs and reports are available in our paper, at https://osf.io/2kdb9/

Regarding **accuracy**, instead, we didn't find meaningful differences between the different shapes. Therefore, we can't say that participants answered more precisely to questions on a shape rather than on another one. We just found out that the **compare** task, that was more memory-intensive than the others, produced worse results.

At the end of the experiment, participants could fill out an optional survey, in which they were asked to express their **preferences** and **opinions**.

We received some interesting comments that pertained directly to the shape of the timelines. This gave us a notion of how the participants felt about certain shapes to an extent that they cared enough to express it in an optional feedback section. In this case too the majority of the comments were in favor of the horizontal and vertical lines. The comments were of the similar trend:

> *"…the vertical line is easy to read"*
> *"The horizontal and vertical graphs were easiest"*

A few remarked in favor of the circular shape:

> "*…but circles were easiest for my eyes*".

The spiral shape also had a few admirers under its belt with comments like:

> *"The planting dataset with the spiral is a good example of a timeline that is functional and visually appealing".*

However, on the flip side, some of these participants revealed concerns about the spiral shape, stating it to be difficult and

confusing. Their comments echoed similarly:

> "*I found the spiral hard to work with*"
> "*…the spirals were confusing*"
> "*The spiral one was really hard*"

Percentage of participants who selected specific shapes as easiest to read for each task.

Considering the findings of our experiment, we formulated some design recommendations for timelines using one of the data set types we took into account. Here is a list of recommendations regarding **timeline readability**:

1. Use linear vertical timelines for situations which require fast data lookup.
2. Avoid spiral timelines when the task requires fast lookup.
3. If you use a more creative, expressive shape, such as a spiral timeline, also include a tutorial or visual cues to assist the user in learning and understanding.

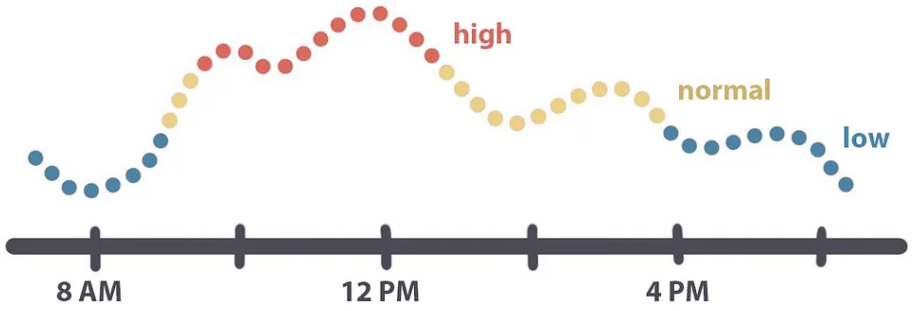
Regarding **task performance**:

- Tasks which require long-term memory dependence like the Compare task are slower and less accurate on all the timeline designs. This effect is amplified with spiral timelines. We recommend using linear timelines for difficult tasks which require complicated decision making.

Regarding **dataset flexibility**:

- Within this study, we did not find that dataset choice affects the readability of the timeline. Therefore, we recommend designers to be flexible with their choice of timeline shape to maximize readability or improve engagement. However, if the dataset is complex, even for mixed data,we recommend using a linear timeline.

We want to stress that the importance of the following points, especially regarding timing, heavily depends on the context. For instance, in emergency medicine a difference of a few seconds may be much more relevant than for reading a timeline published in a magazine.



A timeline representing blood glucose values of a person with diabetes throughout the day. A medical professional may have to look through many of these charts in the span of a short amount of time, therefore this is a situation in which fast data lookup is relevant.

This type of spiral visualization is widely used in schoolbooks to represent how long and complex the history of life on earth has been before humankind. Here, the spiral is used to convey the sense that the more we go back in time, the more our knowledge becomes sparse.

It is also very important to keep in mind the expressiveness of the data — how well the timeline shape can represent the underlying data. For instance, representing linear data on a circular shape may be misleading to the readers by inducing them into thinking there are recurrences in the data when there are none.

We therefore suggest designers carefully evaluate our recommendations based on their domain goals and specific use case.

Timeline visualizations are pervasive in our everyday life. However, we know little about ways to design timelines effectively. In this work, we learned that the shape does indeed affect the readability of timelines. We found good evidence that lines support reading the timelines quicker than the non-linear shapes. We also found evidence that spirals are not only perceived confusing by users but are also slower in lookup tasks. We hope our work will encourage future research on the topic.

Our paper, "*Evaluating the Effect of Timeline Shape on Visualization Task Performance",* written by Sara Di Bartolomeo, Aditeya Pandey, Aristotelis Leventidis, David Saffo, Uzma Haque Syeda, Elin Carstensdottir, Magy Seif-El Nasr, Michelle A. Borkin and Cody Dunne, was accepted at CHI 2020 and can be found at https://osf.io/2kdb9/

Our supplemental material, code for running the experiment and for drawing the timelines, and all of our result data, can be found at: https://osf.io/qr5yu/