Map or Gantt? Which Diagram Helps Viewers Best in Spatio-Temporal Data Exploration Tasks?

Bosveld-de Smet, Leonie; Houben, Daniel

Published in:
Diagrammatic Representation and Inference 11th International Conference, Diagrams 2020, Tallinn, Estonia, August 24–28, 2020, Proceedings

DOI:
10.1007/978-3-030-54249-8_28

Publication date:
2020

Map or Gantt? Which Diagram Helps Viewers Best in Spatio-Temporal Data Exploration Tasks?

Leonie Bosveld-de Smet(✉) and Daniël Houben

University of Groningen, Groningen, The Netherlands
l.m.bosveld@rug.nl

Abstract. In this paper we investigate the effectiveness and efficiency of two two-dimensional static visual representations of spatio-temporal data, a map-based and a Gantt-based diagram, in their support of various information retrieval tasks. The map-based diagram is characterized by a natural spatial arrangement of locations on a schematic map. The Gantt-based one represents time naturally as a linearly ordered set of time intervals from left to right. A within-subject empirical experiment has been conducted, in which participants were asked to verify queries about persons, locations, and time intervals. The formulation of the queries was based on (i) Bertin's three reading levels, (ii) certain cognitive operations, and (iii) different syntactic orders of expressions denoting persons, locations and times. Response correctness and response time were recorded. With respect to response accuracy, both diagrams support viewers well in nearly all information retrieval tasks. Regarding efficiency, the map-based diagram elicited significantly faster response times than the Gantt-based one, except for queries with time in focus. The results suggest that map-based diagrams require less search and reasoning effort of viewers to retrieve the information asked for in the task types used in this study.

Keywords: Information visualization · Spatio-temporal data · Map · Gantt chart · Data exploration tasks · Reading levels · Cognitive operations

1 Introduction

This paper is about visual representations of spatio-temporal data based on two popular static two-dimensional diagrams, the geographic map and the Gantt chart. It is also about data exploration tasks to be performed with the aid of these two diagram types. Both diagram types are adapted to make them suited to the representation of the three components involved in spatio-temporal data, which are objects, locations and times. The paper has two aims. The first is to investigate to what extent the adaptations of the map and the Gantt chart facilitate the performance of certain information retrieval tasks. The second aim is to explore in what way task type and complexity of the information depicted affect task performance.

A.-V. Pietarinen et al. (Eds.): Diagrams 2020, LNAI 12169, pp. 357–364, 2020.
https://doi.org/10.1007/978-3-030-54249-8_28

Spatio-temporal data are data that relate to both space and time. Andrienko *et al.* classify these data according to the kind of changes occurring over time [1]. In this study, the spatio-temporal data are based on appearances of individuals, also called agents, at a certain location in a certain time interval. The pattern of appearance and disappearance of agents at certain locations may be thought of as a continuous path through space-time. The changing states can be fully known, only partially known, or unknown.

Spatio-temporal data are best represented in a way that conforms to human conceptualizations of the world in space and time [7]. For the representation of space, static maps are a popular and intuitive way to show all kinds of geographical information. Map representations use horizontal and vertical axes to encode the canonical world directions of north-south and east-west [8]. Mapping time to space is less obvious. Time is an abstract notion. Its dynamic character suggests a representation by time, rather than space. Dynamic and interactive visual representations of spatio-temporal data can add a dimension to the interpretation of geographic data. Dynamic maps allow to view temporal characteristics of change in 'world time' or 'display time' [3]. Nevertheless, we often use space to reason about time [5]. Timelines are powerful and well-understood metaphors for visualizing time. For most European people, time is naturally represented as a timeline from left to right, or from top to bottom, indicating progression of time [9]. Such a timeline is a linearly ordered set of time points or time intervals within a certain timespan. Calendars, diaries, appointment books and the like make use of this conceptualization of time.

The two diagrams used in this study are adaptations of two popular diagrams most people are familiar with, namely the geographic map and the Gantt chart. Originally neither of these diagram types are designed for the representation of spatio-temporal data. For this study, they are adapted in order to represent all three interrelated components involved in spatio-temporal data, following Kriglstein *et al.* [6]. Figures 1 and 2 illustrate how we have adapted the map and the Gantt chart. In the map adaptation, locations are visualized as circles corresponding to their geographic location. The Gantt adaptation uses the original time representation of the Gantt, as a timeline of time intervals on the x-axis at the top, from left to right, as time proceeds. In the map, time intervals are indicated by numeric annotations next to circle portions. Moreover, residence time corresponds to sizes of circle portions that are calculated relatively to the other agents' residence times at that location. In the Gantt adaptation, locations are placed on the y-axis as textual row labels, from top to bottom, in alphabetical order. Shape and color coding are used for the representation of agents. A legend associates colors to agents' names. The map adaptation represents agents as circle portions in different colors, the Gantt adaptation consists in replacing the bars that originally illustrate projects, by agents' appearances, in different colors. Note that in both adaptations, only one component gets an intuitive representation, which is location in the map and time in the Gantt, while the other components' representations are less intuitive.

For this study, we decided to differentiate agents' paths through time and space according to complexity. We consider a path and its representation as simple in case it implies at most two location changes of an agent. For complex paths the number of an agent's moves varies from two to six. Note that the complexity of a path affects its representation.

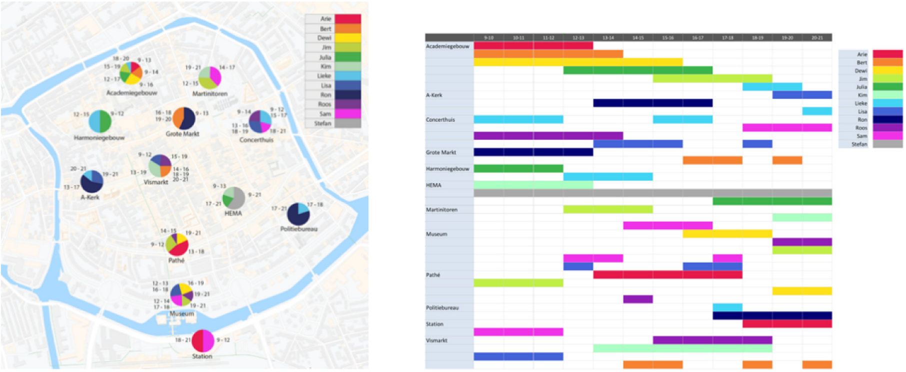

Fig. 1. Map (left) and Gantt (right) depicting the same complex scenario

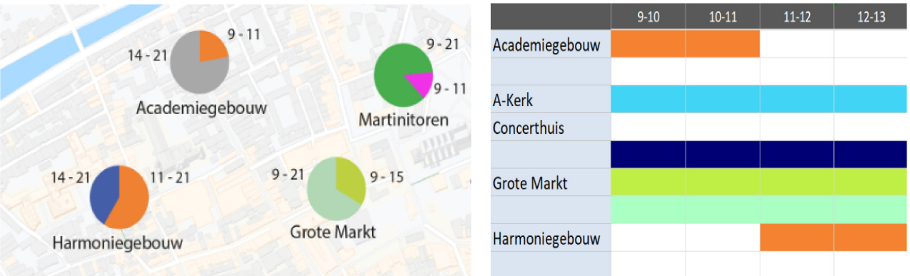

Fig. 2. Parts of Map (left) and Gantt (right) depicting simple scenarios

For the measurement of users' performance with the map-based and the Gantt-based adaptations, we have used a variety of visual data exploration tasks. There are numerous typologies of data exploration tasks [1]. One such typology is suggested by question types and reading levels, proposed by Bertin for arbitrary data [2]. There are as many question types as there are components in the data. For each question type, Bertin introduces three reading levels: elementary, intermediate, and overall. The level of reading indicates whether a question refers to a single component (elementary), to a group of components (intermediate), or to the whole phenomenon characterized by all components together (overall). Peuquet confines Bertin's notions of question type and reading levels to spatio-temporal data [7]. The notion of reading levels can be independently applied to the spatial and to the temporal dimensions of spatio-temporal data. Combinations are possible. Bertin's scheme is not fully satisfying for spatio-temporal data. While it makes explicit that exploration tasks involve identification of (sets of) single elements, it leaves implicit whether other cognitive operations are involved. Within the same question type and reading level, the exploration task may indeed require the analyst to compare or relate two or more (sets of) elements. The identification-comparison dimension should be added to the typology of data exploration tasks [1]. Comparison should be interpreted here in Blok's broad sense of determining relationships [3]. The concrete linguistic specification of data exploration tasks adds another differentiating aspect. Kessell and

Tversky, in their study of matrix-based visualizations of spatio-temporal data, chose to operationalize question types as query statements with different foci on location, time, or person [4]. Queries are categorized according to the element(s) in focus and sentence order. For the purpose of our research, we used a typology of data exploration tasks merging Bertin's reading levels for arbitrary data, Peuquet's specialization of these for spatio-temporal data, Andrienko et al.'s extension of identification with comparison, and Kessell and Tversky's approach to formulate data exploration tasks as query statements with referential expressions of different components as topic. In principle, multiple exploration task types in relation to spatio-temporal data are possible. We selected a restricted set from these, which are categorized and illustrated in Table 1.

Table 1. Query types used in the experiment and examples of query statements of each query type, translated from Dutch.

Reading level and cognitive operation(s)	1st position (component in focus)	2nd position	3rd position
Elementary; Identification	1 agent name	1 time interval (1 hour)	1 location name
	e.g.: **Julia** is from **11 to 12** at **the Pathé**		
Intermediate; Identification	1 time interval (1 hour)	>1 agent name	1 location name
	e.g.: From **13 to 14 Arie and Stefan** are on **the Martinitoren**		
Intermediate; Identification, Comparison and Counting	1 time interval (1 hour)	cardinality of group of agents	1 location name
	e.g.: From **15 to 16** there are **exactly three persons** at **the Vismarkt**		
Overall; Identification and Comparison	1 agent name	whole timeline	>1 location name
	e.g.: **Bert** visits **today the HEMA and the railway station**		
Overall; Identification, Comparison, and Counting	1 location name	whole timeline	cardinality of group of agents
	e.g.: **The A-Kerk** is visited **today** by **exactly five persons**		

In order to get more insight into the information read-off afforded by adaptations of the map and the Gantt chart, and its possible relation with query type and complexity, we have formulated the following three interrelated research questions:

RQ1: To what extent do two specific adaptations of the map and the Gantt chart facilitate naive users in the performance of spatio-temporal data exploration tasks?

RQ2: Does task type, characterized on the basis of (i) reading level, (ii) cognitive operations involved, and (iii) task wording, affect task performance?

RQ3: Does information complexity, based on number of moves of agents through space and time, affect task performance?

2 Method

To answer the research questions, we conducted an empirical experiment with a within-subject design[1]. Forty individuals with different age, gender, and educational background were recruited to take part in the experiment. The sample consisted of 18 male and 22 female participants, with a wide age range from 16 to 78. All participants were native speakers of Dutch, with normal reading proficiency. None of the participants was color blind. All participants were familiar with maps and tabular representations of timelines. Most were more or less familiar with the city center used in the scenarios represented by the diagrams.

All participants were subjected to 32 different queries in total, corresponding to the five different types of data exploration tasks as given in Table 1. Sixteen queries (4 categorized as elementary, 4 as intermediate, and 8 as overall) had to be answered with the map, 8 with the simple version, and 8 with the complex one. The same distribution of queries was used for the Gantt. The visualization types and the queries were presented to the participants in random sequence. Answer accuracy and answer completion time were recorded. Half of the queries are true, the other half false.

The map-based and the Gantt-based visualizations each represent different scenarios of people residing at different locations during certain time intervals. The locations are chosen in the center of the city of Groningen, in the north of the Netherlands. The spatio-temporal data consist of 12 different agents, 12 one-hour time intervals, ranging from 9 A.M. to 9 P.M. and 12 well distinguishable locations (church, shop, cinema, etc.) in the city of Groningen. The movement of agents differs from no movement to a maximum of six location changes. These data are visualized in the map and the Gantt format as shown in Fig. 1 for a complex scenario.

A digital survey was created in Qualtrics. The total survey consisted of (i) an introductory page, explaining shortly the experiment, (ii) the actual survey with the 32 query statements and accompanying visual representations, and (iii) post-task questions, consisting of demographic questions and questions about preference, and insightfulness and aesthetics of the diagram types used in the study.

Each participant took individually part in the experiment in a quiet environment, and used a mouse, keyboard, and display with a minimum resolution of 1600 $\times$ 900 px. The researcher was present in the room, with view on the display and the participant, and took notes, if there was a reason to do so.

[1] The materials will be available for the interested readers. Contact the authors for more information.

3 Results

There was one outlier. The results discussed below are based on 39 × 32 result data points for accuracy and for completion time, each. No effect of age, gender, educational background and preference was found. For the inferential statistics, we have used paired-samples t-tests. An overview of all response accuracy and response time results is shown in Table 2.

Table 2. Mean response accuracy (left) and mean response time (right) results per task type, diagram type, and diagram complexity degree.

	Response accuracy measures (Mean accuracy in percentages)				Completion time measures (Mean completion time in seconds)			
Data Exploration task	Simple versions		Complex versions		Simple versions		Complex versions	
	%	*p*	*%*	*p*	*Mean*	*p*	*Mean*	*p*
Elem. ident.								
Map	92.3	0.744	93.6	0.534	**34**	**<0.001***	**41**	**<0.001***
Gantt	91.1		91.1		**43**		**53**	
Interm. ident.								
Map	87.2	0.711	89.7	0.711	27	0.294	29	0.494
Gantt	89.7		87.2		29		27	
Interm. count.								
Map	92.3	0.570	**94.9**	**0.002***	20	0.901	28	0.918
Gantt	94.9		**64.1**		20		28	
Overall comp.								
Map	98.7	0.999	98.7	0.570	**36**	**0.005***	**38**	**<0.001***
Gantt	98.7		97.4		**42**		**47**	
Overall count.								
Map	98.7	0.103	98.7	0.999	**23**	**0.002***	**21**	**0.002***
Gantt	93.6		98.7		**30**		**40**	

Results in **Bold**, and p-values in **Bold** with asterisk indicate significant Map-Gantt differences.

Both map-based and Gantt-based representations of spatio-temporal data are nearly equally effective. Regarding overall response accuracy, there is no significant difference between the map and the Gantt ($t(38) = -1.859$, $p = 0.0707$ (two-tail)). Neither task type, nor complexity seems to affect accuracy performance. The only significant difference in accuracy performance we have detected is one between the complex map and the complex Gantt, for the intermediate search level task involving identification of a single time interval and a single location, and counting of number of agents visiting that location at that time. This task turned out to be more difficult to perform with the aid of the complex Gantt (only 24 of the 39 queries, 64.1%, were answered correctly) than for the complex map (37 correct answers out of 39, 94.9%). At α-level 0.05, this difference is significant ($t(38) = -3.376$, $p = 0.002$ (two-tail)), in favor of the map.

The efficiency results clearly show that the map-based representations are more facilitative than the Gantt-based ones. On average the participants took six minutes to verify the 16 queries with the Gantt, and took five minutes with the map to verify 16 similar queries. Overall, the participants performed the tasks significantly faster with the map than with the Gantt ($t(38) = 7.223$, $p = <0.001$ (two-tail)). The completion times per task type given in Table 2 shows much more diversity than the response accuracy results. Significant differences have been found for tasks with a component in focus other than time. The participants performed the elementary identification tasks significantly faster with both the simple and complex versions of the map (simple: $t(38) = 4.566$, $p = <0.001$ (two-tail); complex: $t(38) = 4.580$, $p = <0.001$ (two-tail)). Similar results are observed for the overall search tasks, with either agent or location in focus, and involving only comparison, or comparison and counting. Both simple and complex versions of the Gantt turned out to be significantly less supportive than their map variants (overall comp. simple: $t(38) = 2.958$, $p = 0.005$ (two-tail); overall comp. complex: $t(38) = 4.793$, $p = <0.001$ (two-tail); overall count. simple: $t(38) = 3.361$, $p = 0.002$ (two-tail); overall count. complex: $t(38) = 3.286$, $p = 0.002$ (two-tail)). We observe that complexity of scenario depicted tends to increase response times. Yet, this increase is not significant.

4 Discussion

Both map-based and Gantt-based adaptations allow naive users to answer the query statements quite accurately, without there being a significant difference between the two diagram types. There is only one exception. The participants made significantly more errors with the Gantt-based representation depicting complex information in their performance of intermediate reading level tasks, involving identification, comparison, and counting, with focus on time in wording. We suspect that this result is due to the visual delimitation of location. In the map, the locations are clearly delimited by their circular shape. The Gantt cells don't have clear contour lines for location. During the experiment some participants used their finger on the screen to make sure whether a bar was part of the row corresponding to some location. The map-based adaptation is, however, significantly more facilitative with respect to efficiency, especially in tasks which do not have time in focus (**RQ1**).

Task type has barely influence on effectiveness, but it does affect efficiency. Focus on time in task wording seems to lead to relatively faster response times in the Gantt-based diagram, but does not seem to be harmful for the map-based one, although time does not get an intuitive representation here. Task wording has an effect, but this effect seems to be rather moderate, as it did not help participants to be more accurate in the complex case (**RQ2**).

The results do not suggest any clear effects of complexity of information depicted in the diagrams, neither on effectivity, nor on efficiency of task performance. Only in one specific case participants performed significantly worse, but in this case the participants were relatively faster, due to the wording of the task. So here again, the effect is moderate. A visually clearer design of location rows in the Gantt might eliminate this significant difference (**RQ3**).

In conclusion, in their specific forms chosen for this study, maps fit cognitively better to spatio-temporal exploration tasks than Gantt charts do, at least with respect to efficient scanning. There is a limited effect of task type, and a very modest effect of complexity of information represented on task performance. Understanding of and reasoning with diagrams involve an interaction between a variety of factors: (i) a diagram's inherent structural properties, (ii) specific design options chosen, (iii) match between task requirements and ease of information retrieval afforded by the diagram, and (iv) individual user's characteristics such as age, prior knowledge, cognitive style, and preference. Inherent structural properties of the two diagram types in combination with the design options used may be responsible for the better performance of the map. Also, more familiarity with a map than with a Gantt chart, and prior knowledge of the city center depicted may have worked in favor of the map. Based on the results of this study, involving only a restricted set of task types, we are not able to separate the wheat from the chaff. The results of the current study do not provide compelling evidence about which factor has been most influential. Further research has to be performed.

References

1. Andrienko, N., Andrienko, G., Gatalsky, P.: Exploratory spatio-temporal visualization: an analytical review. J. Vis. Lang. Comput. **14**(6), 503–541 (2003)
2. Bertin, J. Sémiologie graphique: les diagrammes. Mouton, Paris (1967)
3. Blok, C.: Monitoring change: characteristics of dynamic geo-spatial phenomena for visual exploration. In: Freksa, C., Habel, C., Brauer, W., Wender, Karl F. (eds.) Spatial Cognition II. LNCS (LNAI), vol. 1849, pp. 16–30. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-45460-8_2
4. Kessell, A., Tversky, B.: Visualizing space, time, and agents: production, performance, and preference. Cogn. Process. **12**(1), 43–52 (2011)
5. Kriglstein, S., Pohl, M., Smuc, M.: Pep up your time machine: recommendations for the design of information visualizations of time-dependent data. In: Huang, W. (ed.) Handbook of Human Centric Visualization, pp. 203–225. Springer, New York (2014). https://doi.org/10.1007/978-1-4614-7485-2_8
6. Kriglstein, S., Haider, J., Wallner, G., Pohl, M.: Who, where, when and with whom? Evaluation of group meeting visualizations. In: Jamnik, M., Uesaka, Y., Elzer Schwartz, S. (eds.) Diagrams 2016. LNCS (LNAI), vol. 9781, pp. 235–249. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-42333-3_19
7. Peuquet, D.J.: It's about time: A conceptual framework for the representation of temporal dynamics in geographic information systems. Ann. Assoc. Am. Geogr. **84**(3), 441–461 (1994)
8. Tversky, B.: Distortions in memory for maps. Cogn. Psychol. **13**(3), 407–433 (1981)
9. Tversky, B., Kugelmass, S., Winter, A.: Cross-cultural and developmental trends in graphic productions. Cogn. Psychol. **23**(4), 515–557 (1991)