A Comparison of Radial and Linear Charts for Visualizing Daily Patterns

Manuela Waldner, *Member, IEEE CS*, Alexandra Diehl, Denis Gračanin, *Senior Member, IEEE*,
Rainer Splechtna, *Member, IEEE CS*, Claudio Delrieux, and Krešimir Matković, *Member, IEEE CS*

Abstract— Radial charts are generally considered less effective than linear charts. Perhaps the only exception is in visualizing periodical time-dependent data, which is believed to be naturally supported by the radial layout. It has been demonstrated that the drawbacks of radial charts outweigh the benefits of this natural mapping. Visualization of daily patterns, as a special case, has not been systematically evaluated using radial charts. In contrast to yearly or weekly recurrent trends, the analysis of daily patterns on a radial chart may benefit from our trained skill on reading radial clocks that are ubiquitous in our culture. In a crowd-sourced experiment with 92 non-expert users, we evaluated the accuracy, efficiency, and subjective ratings of radial and linear charts for visualizing daily traffic accident patterns. We systematically compared juxtaposed 12-hours variants and single 24-hours variants for both layouts in four low-level tasks and one high-level interpretation task. Our results show that over all tasks, the most elementary 24-hours linear bar chart is most accurate and efficient and is also preferred by the users. This provides strong evidence for the use of linear layouts – even for visualizing periodical daily patterns.

Index Terms—Radial charts, time series series data, daily patterns, crowd-sourced experiment

---

◆

---

1 INTRODUCTION

Rose charts, also called coxcomb charts or polar-area charts, became popular in the 19th century when Florence Nightingale visualized mortality in the Crimean War from April 1854 to March 1855 [10, 16]. The charts show the number of soldier deaths for each month split into two rose charts. Since then, rose charts or similar radial charts have been frequently used to depict periodical temporal patterns, such as frequencies of daily news reports [45], 24-hours network traffic [22, 37], distribution of phone calls over the day [46], daily traffic information [34], climate effects over a year [44], or monthly sales [42]. It is often argued that mapping periodical behaviors to a radial chart is an intuitive and appropriate representation of the underlying data [10, 12, 37].

However, there is empirical evidence that radial charts are inferior for many analytic tasks as compared to linear ones [6, 12, 15, 29, 31, 47]. This has even been shown for radial charts encoding periodical time-dependent data, where linear charts outperformed or performed equally well as radial ones when conducting low-level analytic tasks [1, 11]. These studies, however, compared weekly, monthly, or yearly periodical patterns. To the best of our knowledge, Fuchs et al. [26] presented the only user study comparing different glyph layouts for monitoring tasks of *daily* patterns. They found that radial glyphs are more effective for reading values at specified time points. In particular, their users reported to like the clock metaphor of the color-coded 24-hours clock glyph. However, such a glyph does not utilize the layout of an analog clock. For future work, some of their users therefore suggested to visualize twelve hours per glyph to create a more intuitive mapping.

Hours of a day are a special domain, as observers are trained to read the hours from radial clocks from an early age [9, 25]. It is therefore possible that familiarity with the radial clock design facilitates a more efficient cognition of daily patterns in the underlying data when visualizing the data with a radial clock-like layout. This is particularly relevant for non-expert audiences, where the goal is to communicate information of general interest, such as daily traffic information or social media presence over the day, effectively, but also in an engaging manner [39].

We present the first controlled user study investigating whether imitating a radial analog clock layout can facilitate the interpretation of *daily* patterns. Our main hypothesis was that visualizations of daily patterns using a clock-like radial layout can be interpreted more efficiently than using alternative radial or linear charts by untrained users due to the familiar layout. We conducted a crowd-sourced experiment where we compared the users' performance with two layouts (linear and radial) with two different encoding cardinalities (2x12 hours and 1x24 hours). Thereby, the two juxtaposed radial rose charts use the same encoding principle as a radial clock, where the twelve hours of a.m. and p.m. are arranged in one circle each. The single 24-hours radial variant, where noon is located at the bottom and midnight at the top, can also be observed in many scientific works (e.g., [24, 34, 40, 44–46]) and in-the-wild (e.g., [35]), but does not correspond to the familiar 12-hours clock representation. We contribute evidence that a clock-like radial layout does not have a positive impact on understanding daily patterns for untrained users. We also provide new findings about the effect of chart separation on users' task performance for both linear and radial charts. Finally, we performed an exploratory analysis of error cases and report observations that can explain why untrained users often misinterpret rose charts.

2 BACKGROUND

Radial charts are often considered more aesthetic and space-efficient [12, 32, 37], as well as more natural and therefore more memorable [8] than linear charts. They are commonly used to visualize hierarchical data [49], multivariate data [33, 54], or set relations [3]. Draper et al. [20] provide a survey of radial charts up to 2009. Of our special interest is their presumable intuitiveness for visualizing time-dependent periodical data [10, 12, 37].

Rose charts [10, 16] can be seen as radial histograms of time-dependent data, where each bar represents the frequency for a discrete temporal bin, such as one month or one hour. Circular silhouette graphs encode frequencies as continuous curve following a yearly circle [30]. Brehmer et al. [11] visualize value ranges as bars, both in a linear and a radial layout. Alternatively to bars, values can be represented by lines in polar coordinates [22]. The peaks of these star glyphs can addition-

---

- *Manuela Waldner is with TU Wien. E-mail: waldner@cg.tuwien.ac.at.*
- *Alexandra Diehl is with University of Zurich. E-mail: diehl@ifi.uzh.ch.*
- *Denis Gračanin is with Virginia Tech. E-mail: gracanin@vt.edu.*
- *Rainer Splechtna is with VRVis Research Center. E-mail: Splechtna@VRVis.at.*
- *Claudio Delrieux is with Electric and Computer Eng. Dept., Universidad Nacional del SUR and CONICET. E-mail: cad@uns.edu.ar.*
- *Krešimir Matković is with VRVis Research Center. E-mail: Matkovic@VRVis.at.*

ally be connected by lines [27]. Finally, time-dependent values can be encoded by color resulting in clock glyphs [26] or "Kaleidomaps" [4].

An alternative to juxtaposition of multiple radial graphs to show longer time sequences, like the two years in Nightingale's rose charts [10, 16], is to arrange longer time sequences as concentric rings [14, 21] or in a spiral [19]. This principle can also be used for visualization of longer time series data, where each year, for instance, is visualized in 360 degrees [51, 53].
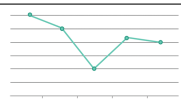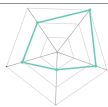
Several studies indicated that radial charts are less effective than linear charts for exploratory tasks, relative judgments, and readability. For instance, it has been shown that relative judgments are more difficult to perform using pie charts than bar charts [15, 23, 31, 47] because area and angle judgments are less accurate than position or length judgments [15]. In addition, non-salient angles (*i.e.*, not $0°$, $90°$, etc.) are hard to interpret in pie charts [23, 47]. For small-scale charts on smart watches, Blascheck et al. [6] showed that classic linear bar charts are also easier to read than donut charts or radial bar charts, where bars are represented as arcs. Goldberg and Helfman [29] compared different variations of linear and radial fixed-sized charts in an eye tracking study. They could show that reading values is clearly more difficult in radial charts because mapping data points to values is more inefficient compared to linear charts. Through a large-scale online study, Diehl et al. [18] showed that memorizing specified locations is also more difficult for radial visualizations than Cartesian visualizations, but circle sectors are easier to memorize than rings. On the other hand side, Spence [48] showed that radial charts perform as well as linear charts for comparing proportions, and that they provide better visual cues for part-whole estimations and natural anchors ($0°$, $90°$ etc.). These studies did not focus on time-dependent data.

Several studies compared linear visualizations for time series data. Javed et al. [36] found that separating multiple time series into small multiples is more effective than superimposing them. Gogolou et al. [28] compared three linear time series visualization techniques (line charts, horizon charts, and color fields) and concluded that the visual encoding has a lot of influence on the perception of similarity between two time series. Albers et al. [2] compared different linear visualization techniques for different tasks analyzing time series data. They found that a composite graph, overlaying a line graph with a bar chart representing the averages of the bar's region, outperformed other time series representations.

Of most relevance for the encoding of daily patterns are studies comparing linear and radial layouts for periodical time series data. Adnan et al. [1] compared different linear and radial layouts for visualizing time series data across several weeks. For all tasks, the linear charts outperformed the radial charts. Brehmer et al. [11] systematically compared linear and radial bar charts for visualizing weekly, monthly, and yearly patterns on mobile phones. They found that linear charts were generally more efficient, while the accuracy was comparable to radial charts. Our study differs as we look into the special case of daily patterns that may benefit from the visualization in a clock-like chart.

To the best of our knowledge, Fuchs et al. [26] presented the only study comparing different glyph designs, including linear and radial variants, for *daily* patterns in different data granularities from 24 to 96. They systematically compared four different glyphs: radial star and color-coded clock glyphs and linear line and colored stripe glyphs. They found that linear glyphs were most accurate and efficient for value comparison, but the radial glyphs were more accurate and efficient, respectively, for time comparisons. Informal feedback by their student participants suggests that the clock glyph was appreciated for its obvious clock metaphor. This inspired us to investigate this aspect in more detail by systematically varying both, the layout (linear and radial) and the encoding granularity, to test the effect of the clock design on the task performance with large-scale visualizations rather than glyphs. As we were interested in the appropriateness of clock-like radial charts for showing daily patterns for a broad audience, a difference of our study is that we performed a crowd-sourced study with potentially untrained users. The population registered at crowd-sourcing platforms is expected to differ from student subjects in multiple aspects, such as a wider age range and higher openness (imagination, curiosity) [38].

Table 1: Design space of linear and radial visualizations for periodical time-dependent patterns, organized by encoding (rows) and layout (columns). In our study, we investigated bar encodings (second row).

3 VISUALIZATION OF DAILY PATTERNS

The variety of techniques to visualize periodical time-dependent patterns in Section 2 shows that there is a large possible design space. This design space includes linear charts, such as line charts, bar charts, or colorfields, as well as radial charts, such as star glyphs, rose charts, or color-based clock glyphs (Table 1). In this design space, we only consider radial charts where the time is encoded in angle and the value in radius or color. Radial bar charts sometime also use angle to encode the value (e.g., [6]), but we do not consider this case here.

Our goal was to visualize aggregate values across 24 hours, such as the accumulated energy consumption per hour or the frequency of traffic accidents for each hour a day. This corresponds to a periodical temporal histogram with 24 bins. We excluded color as it is expected to be inferior to position encodings [2]. Goldberg and Helfman [29] showed that across linear and radial charts, data points could be found more efficiently using bar charts than line charts. Also, as we are dealing with binned data, interpolating data points with line connections could lead to misinterpretations. According to Borkin et al. [8], bars – including linear bar charts, stacked bar charts, and circular bar charts – are also the most commonly used encoding in news media and government reports. We therefore chose a classic bar histogram for visualizing our data. In polar coordinates, this corresponds to a rose chart.

This encoding choice leaves us with four simple visualizations for 24-hours time series data: two juxtaposed 12-hours radial charts (**12r**), a single 24-hours radial chart (**24r**), two juxtaposed 12-hours linear bar charts (**12l**), and a single 24-hours linear bar chart (**24l**). The 12-hours radial chart (12r) adopts the ubiquitous clock layout with twelve on the top and six at the bottom. At the same time, it suffers from the known disadvantages of radial charts, such as the difficulty to decode values accurately and efficiently. In addition, to encode the 24 hours of the day, the data has to be split into two separate charts. This means that, in contrast to the 24-hours radial chart, the data cannot be read without switching charts. This is also true for the two 12-hours linear bar charts (12l). However, in this case, the right-most bar of the first 12-hours chart marks a clear end of the a.m. section. Continuing to the beginning of the p.m. section is supported by the natural reading order. On the other hand, the separation into a.m. and p.m. may also be advantageous as the visualized data is separated into smaller chunks with two identical value axes. The 24-hours radial chart (24r) is the only visualization that is truly continuous so that also temporal patterns across midnight are not separated.

4 EXPERIMENT

The goal of this experiment was to verify our hypothesis that daily patterns can be interpreted most efficiently when visualized with a clock-like radial layout. We also hypothesized that for a more detailed analysis, like reading or comparing values, a linear layout will be superior, as suggested by prior work [1, 11, 29]. Therefore, it is sensible to hypothesize that there is no benefit for a 24-hours radial chart, as it

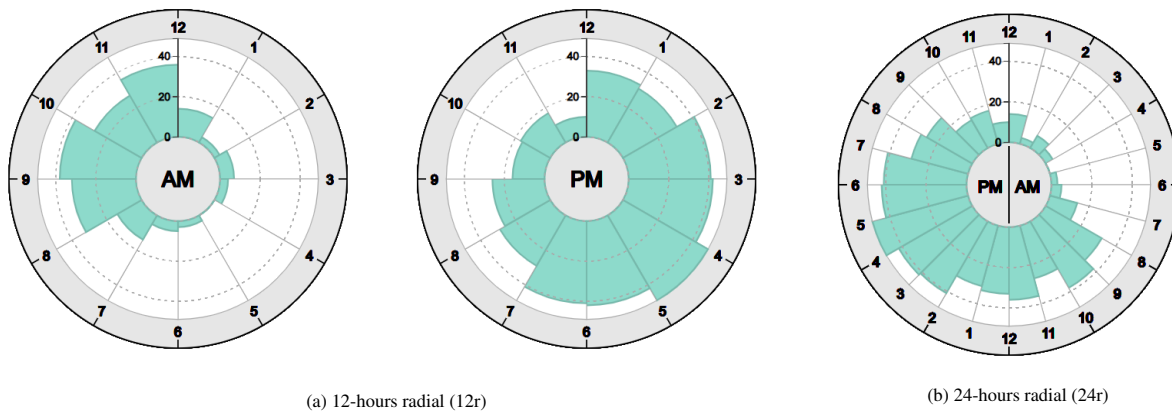(a) 12-hours radial (12r)　　　　　　　　(b) 24-hours radial (24r)

Fig. 1: 2x12-hours and 24-hours radial charts showing the number of traffic accidents in Staten Island, New York, in March.

suffers from known decoding deficiencies, and at the same time does not provide the clock metaphor benefit of the 12-hours radial chart. As our goal was to investigate how well visual encodings are understood and accepted by a general audience, we crowd-sourced our experiment using Amazon's Mechanical Turk system, commonly used to perform graphical perception studies [7, 31, 38]. Also, previous studies comparing time series visualizations have been conducted remotely through Mechanical Turk [2, 11].

4.1　Visualizations

To test these hypotheses, we created four visualizations: two juxtaposed 12-hours radial charts (12r, Fig. 1a), a single 24-hours radial chart (24r, Fig. 1b), two juxtaposed 12-hours linear bar charts (12l, Fig. 2), and a single 24-hours linear bar chart (24l, Fig. 3). These visualizations differ in two aspects: their encoding **cardinality**, which can be 12 hours or 24 hours, and **layout**, which is radial or linear. We recruited our study participants from North America. Therefore, we encoded time in the 12-hours clock system using AM and PM labels.

As our focus was to investigate visualizations for the masses, we used a constant chart size, representative of visualizations found in news or on websites. Borkin et al. [8] used up to 512×512 pixels for their representative in-the-wild visualizations in the study. Similarly, we rendered the two juxtaposed 12-hours radial charts with a radius of 200 pixels. We left a 60 pixel radius inner circle, where we placed the AM and PM labels. The outer ring, where the hour labels are located, has a darker color and uses additional tick marks to emphasize the clock layout. Like the original Nightingale rose chart, the two 12-hours radial charts were rendered next to each other to make the best use of conventional monitor aspect ratios. On small screens, they were rendered on top of each other.

The 24-hours radial chart is rendered with the same radius. This makes the bars narrower compared to the two 12-hours charts. Like for the 12-hours variant, the bar value is encoded along the radial axis. We decided to keep the radius constant to keep the value encoding constant across the conditions. This means that the 24-hours radial chart is more space efficient compared to the 12-hours variant. In the inner circle, we placed the label "PM |AM" for the two circle halves.

The linear bar charts (12l, 24l) were rendered with a height of 140 pixels. In this way, the y-axis scale is the same size as the scale in the radial charts. The width of the bar charts is the same as the circumference of the circle, which corresponds to the medium value grid line in the radial charts. The two 12-hours linear bar charts were rendered with the same width as the 24-hours bar chart to keep the difference between 12-hours and 24-hours variants consistent between radial and linear layouts.

Both linear bar charts have AM and PM labels centered underneath the hour label "6". The 12-hours bar charts have two identical y-axes for both charts. In contrast to the 12-hours radial charts, we always rendered the two 12-hours bar charts on top of each other for two
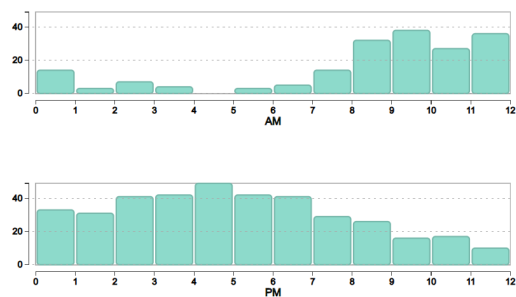

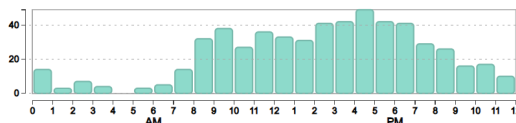Fig. 2: 2x12-hours linear bar charts (12l).

Fig. 3: 24-hours linear bar charts (24l).

reasons: First, the difference to the 24-hours linear chart is minimal when rendering the two charts next to each other. Second, with a vertical alignment, AM and PM values are placed directly on top of each other for easier comparison.

4.2　Data

As our experiment also included a high-level analysis task (Section 4.3.1), we needed a generally understandable data set. We used New York traffic accident data from 2013 to 2016 [52]. We binned that data into boroughs, months, and hours of the day so that each bin gives the accumulated number of accidents per month and per hour of the day for a selected borough.

For the final experiment, we used Staten Island borough data from 2016, as the number of traffic accidents and their distribution across the time of the day varied most among months (compared to other boroughs). These twelve months have distinct patterns to circumvent the learning effect when performing repeated measures. We did not systematically vary the data as a prior experiment comparing radial and linear charts by Brehmer et al. [11] has already shown that the data characteristics have no influence on the users' performance, as long as the periodic nature of the visualization fits the data (e.g., sleep patterns over the week, or temperature distributions over the year).

4.3 Tasks

We included one high-level analysis task, four low-level analytic tasks, and one subjective rating task to test our hypotheses. The four low-level analytic tasks (tasks 2–5) were adopted from Brehmer et al. [11]. Some of these low-level analytical tasks, such as finding the maximum bar, do not take the temporal component into account. For these tasks, the influence of the clock-like layout therefore can be considered negligible. However, as rose charts are slightly different from the radial bar charts used in the study by Brehmer et al. [11], replicating these results helps to generalize their findings to other types of radial charts. Every task was preceded by a task description before showing the actual visualization. Below, we explain the tasks together with our outcome hypotheses and the measures we obtained to verify the hypotheses.

4.3.1 Task 1: Untargeted Analysis

The purpose of this task was to simulate confronting with the visualization in-the-wild, such as on a news website, without any prior instructions how to read the visualization. Therefore, we did not provide any visualization reading instructions. Before showing the visualization, we only informed the users that we would very briefly show how the number of traffic accidents in Staten Island is distributed during the 24 hours of the day. We asked the users to carefully inspect the visualization during the few seconds that it will be visible, and afterwards to write down everything they observed, even if they are unsure. We showed the non-interactive visualization for ten seconds, after which it was replaced by a text field, where users could take note of their observations. Using this task, our goal was to probe which high-level observations untrained users made within a short period of time. This includes observations of peaks, comparisons of longer periods (such as morning compared to afternoon), or minima. Our hypothesis was that *the 12-hours radial layout (12r) will lead to more observations than any other layout* (**H1**), as the familiarity with the clock layout will facilitate the identification and relation of time periods so that general trends and patterns can be perceived more efficiently.

4.3.2 Task 2: Locate Time

The goal of this task was to assess how quickly and accurately users can find a specified hour of the day in the visualization. We randomly picked a one-hour period (e.g., 8–9 a.m.) and asked the user to click on the bar corresponding to the specified time interval. This one-hour interval was randomly selected from 22 hours of the day. We explicitly excluded the periods between twelve and one to avoid confusion between 12 a.m. and 12 p.m. Our hypothesis was that *selecting a one-hour interval on a 12-hours chart (12r, 12l) is more efficient than on a 24-hours chart* (**H2**). The rationale for this is that separation into a.m. and p.m. would facilitate the search for the target time value compared to a single, continuous encoding. In accordance with user feedback given during a study by Fuchs et al. [26], and because children are usually able to read hour intervals from analog clocks from age five to seven [9, 13, 25], we furthermore hypothesized that *the familiar radial clock layout (12r) will facilitate the location of time intervals compared to the linear layout (12l)* (**H2**). We logged which bar the user selected, as well as the time between visualization onset and bar selection.

4.3.3 Task 3: Read Value

The aim of this task was to read a value for a specified hour. To separate the reading task from the interval location task, the target interval was specified by an arrow. The target interval was selected randomly from all 24 hour intervals. Based on prior empirical evidences [11, 29], our hypothesis was that *reading values is more accurate and efficient using a linear layout (12l, 24l) than a radial layout.* Due to the separation into two bar charts with their own y-axes, we hypothesized that *reading values is most accurate and efficient using the 12-hours linear layout (12l)* (**H3**). We logged the user responses and the time from visualization display to selection of the "continue" button. To quantify the accuracy, we computed the absolute difference between the user response and the actual hour value.

4.3.4 Task 4: Locate Maximum

For this task, we asked users to quickly select the maximum value from the visualization. In contrast to Brehmer et al. [11], we did not query the minimum value as our dataset had a lot of very low or even zero hour values so that this task was not meaningful in our case. Our hypothesis was that *finding the maximum value is most efficient using the 24-hours linear layout (24l)* (**H4**), since in this continuous, linear arrangement, the maximum value clearly stands out. We logged the value of the selected bar and the time until the bar was selected.

4.3.5 Task 5: Compare AM/PM Interval Values

Users were asked to decide as to whether in a one-hour time interval marked by a blue arrow, there were fewer, the same number, or more traffic accidents than in the reference hour marked by a red arrow. The response had to be selected from three radio buttons beneath the visualization. The location of the blue arrow was selected randomly from all 24 hours. The red arrow was then shifted exactly for twelve hours forwards or backwards. This means, we always compared the same time intervals in a.m. and p.m. We made this decision as different intervals may confound with the layout and the cardinality. For instance, we can expect an interaction effect between cardinality and layout for comparing adjacent intervals depending on whether the adjacent intervals are both within a.m. or distributed among a.m. and p.m. In order not to test all these possible arrangements, we decided to stick to this one specific comparison case.

We hypothesized that *any linear layout (12l, 24l) facilitates a more effective comparison of a.m./p.m. interval values* (**H5**). The 24-hours bar chart benefits from the common baseline of the linear bars, while the 12-hours bar charts ensure that common a.m./p.m. intervals are always located on top of each other. We logged the time between visualization display and selection of the "continue" button, the actual value difference, as well as the user's selected radio button. As a correctness measure, we computed if the selected radio button corresponded to the sign of the value differences.

4.3.6 Task 6: Subjective Rating

We asked users to state their subjective assessment for the respective visualization to show traffic accident frequencies on a Likert scale from 1 ("it is not suitable at all") to 5 ("there is no better way to show it"). Our hypothesis was that *users rate the 24-hours radial layout (24r) lowest* (**H6**) as it has no apparent advantages for the presented tasks.

4.4 Design

We used a mixed design with the two independent variables **layout** (radial or linear) as within-subjects factor and encoding **cardinality** (12 or 24) as between-subjects factor. This means, every user performed all six tasks with both, a radial and a linear layout, resulting in 6 (tasks) × 2 (layouts) = 12 trials per user in total. We chose encoding cardinality as between-subjects factor as we believe that switching the cardinality with the same layout during the study may lead to misinterpretations, while switching between linear and radial layout is sufficiently different so that users can readily adjust their decoding strategy. The dependent variables were the number of observations (task 1), accuracy and task completion time (tasks 2-5), and subjective rating (task 6).

The six tasks were always presented in the same order as listed in Section 4.3 – first for the radial layout, then for the linear layout, or vice versa. The presentation order of radial and linear was balanced between the workers. The twelve data sets described in Section 4.2 were assigned to the twelve tasks in random order. With task 1 as the first task for a given layout, we could test whether users would understand the layout and get some first impression of the data without any further guidance – as if they would be confronted with the visualization in-the-wild. For the subsequent tasks, where completion time was a concern, they were therefore already familiar with the encoding.

4.5 Apparatus

Users signed up for a single Human Intelligence Task (HIT) with a link to an external server. We performed a pilot study with five users, a mix of visualization experts and non-experts. The average study completion

time was 15 minutes. We therefore configured the HIT to stay active for 30 minutes. The visualizations were created using the D3 JavaScript library. Users were assigned an anonymous PHP session ID, which was logged together with their responses on the external server. The server script sequentially assigned workers to one of the four possible configurations (cardinality 12 and radial first, etc.).

As most of our tasks were rather simple to solve, we inspected the users' responses to decide whether to accept the HIT for the user. This is a common quality assurance procedure for crowd-sourced studies [7]. Upon completion of the study, the server displayed a code containing the session ID and an estimated error code, which users had to copy to the Mechanical Turk response field. The error code was increased if the text field for task 1 was empty, the selected bar in task 2 was incorrect, the entered value or selected maximum value in task 3 and task 4, respectively, deviated more than 10% from the actual value, and if the rather simple comparison task 5 was answered incorrectly. When users received an error code higher than 2, we manually checked the server logs for their responses and excluded cases that did not provide any meaningful qualitative observations in task 1. Users received a 2 USD compensation if they completed all 12 tasks and the quality criteria above were fulfilled. On average, users required 13 minutes to complete the study.

4.6 Participants

We set 100 users as target for our HIT, which is within the common range of participants in crowd-sourced studies [7] and was also used by the methodologically most similar prior study to ours by Brehmer et al. [11]. We only allowed participants from North America. In the USA, telling and writing time is taught in maths courses from the first grade [17].

From the 100 obtained responses, we rejected eight because they did not provide meaningful responses to task 1 and had more than two incorrect responses in total (see Section 4.5). Age ranged from 19 to 68 with a mean age of 36.2, and 41.3% of the users were female. We did not perform a color-blindness test, since color was not used to encode data. As the labels of the visualizations were larger than the font used for the introduction and task description, we also did not perform any additional visual acuity screening but rather expected anybody able to read the instructions to be a plausible reader of the visualization. Most users reported that they are not at all (19.6%) to averagely (51%) familiar with information visualization. Eight users stated that they are quite familiar, but no user claimed to be a visualization expert.

4.7 Analysis

For the untargeted analysis (H1), we performed a qualitative data analysis using open coding of the obtained text responses with three independent human coders. In the first step, we established a codebook to characterize high-level observations by the users. The coders iteratively adjusted the codebook and discussed conflicting codes until they reached an inter-coder reliability of > 0.8 according to Krippendorff's alpha. The final codebook contained the following codes:

- **Number of observations** identified per response.

- **Comparison** across different points in time or time intervals (e.g., between morning and afternoon).

- Whether the user reported **salient features**, such as peaks.

- Whether the user described points in time and time intervals **quantitatively** (e.g., "8 a.m." or "4p.m. to 5 p.m.").

- Whether the user reported time **qualitatively**, such as "morning" or "late afternoon".

Coders assigned a numeric code for the number of observations and binary values for the remaining four codes to each response. For the statistical analysis, we picked the average number of coded observations by the three independent coders (Fig. 4). For the other codes, we only picked 1 if at least two coders agreed, otherwise 0 (Fig. 5). From

the qualitative time references, the coders extracted ten commonly mentioned day time periods (Fig. 6). Each observation was then assigned to one or multiple time periods.

The low-level analytic tasks (H2 – H5) were evaluated with respect to accuracy by comparing the number of errors, and with respect to efficiency by comparing completion times. Accuracy is visualized as the percentage of overall correct responses in Fig. 7. Error cases were removed for the completion time analysis and qualitatively explored to find potential explanations for the quantitative observations. Results are plotted as means with 95% confidence intervals in Fig. 8.

Finally, we analyzed the obtained subjective ratings on a five-point Likert scale (H6) (Fig. 10). Each optional textual feedback was coded if it only contained negative, only positive, or both types of utterances (*i.e.*, neutral).

All dependent variables were statistically analyzed using analysis of variance (ANOVA) with layout as within-subjects and cardinality as between-subjects factor. Completion times were log-transformed. For the analysis of binary codes in task 1, we used a Generalized Linear Mixed Model (GLMM) using a binary logistic regression model. All F-scores, descriptive statistics, details about the exploratory error case analyses, coded textual feedback, task descriptions, example stimuli etc. can be found in the supplemental study information document.

5 RESULTS

In total, we had 92 users that were accepted for the HIT. Each user performed each task once with linear, once with radial layout, so that we obtained 184 samples in total. 44 users performed cardinality 12, 48 users cardinality 24.

5.1 Task 1: Untargeted Analysis

An ANOVA on the coded numbers of observations showed a medium main effect for layout ($F_{1,90} = 8.852; p = .004; \eta_p^2 = .090$) but not for cardinality ($F_{1,90} = .187; p = .667; \eta_p^2 = .002$), and no interaction between the factors ($F_{1,90} = .458; p = .500, \eta_p^2 = .005$). On average, users reported 1.5 observations using the linear layout and 1.2 observations using the radial layout (Fig. 4).
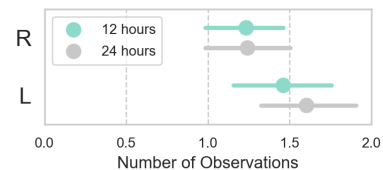
Fig. 4: Mean number of observations with 95% confidence intervals in task 1 for radial (R) and linear (L).

From the 184 coded non-empty responses, 24 (13%) by 16 users did not contain any observations about the data. The responses without any data observations are almost equally distributed among the four conditions. From these 24 responses, 16 were contributed by eight participants, which indicates that they did not understand the task and reported irrelevant or superficial observations in both conditions. The remaining eight were either task misinterpretations that only occurred in the first condition (three cases), such as *"The number of traffic accidents every hour?"* for 12l, or explanations why the user did not come up with any data observations in the second condition (five cases), such as *"the wheel was confusing, I didn't really understand any of it"* for 24r. As the 16 users did not perform worse than the other users in the remaining tasks, we did not exclude them from further analyses.

The number of responses containing comparisons was generally low (16, which corresponds to 7%). There are more responses containing comparisons using the radial layout (12r: 14%, 24r: 12.5%) than using the linear one (12l: 2%, 24l: 6%). However, this difference did not reach statistical significance in our study. Salient features were observed much more frequently (58%), with a slight, non-significant tendency towards 24-hours representations (60% for 24l and 65% for

5

24r) than 12-hours representations (54% for 12l and 52% for 12r). The number of responses that contain quantitative time identifiers range between 56% (24l) to 65% (24r), and no factor was found to have a significant influence on the usage of quantitative reporting of time. However, we found a difference between qualitative time reporting for layout ($F_{1,180} = 8.111; p = .005$): users reported times significantly more often qualitatively using a linear layout (12l: 64%, 24l: 69%) than using a radial layout (12r: 50%, 24r: 56%). The percentage of observations associated with the respective codes is visualized in Fig. 5.
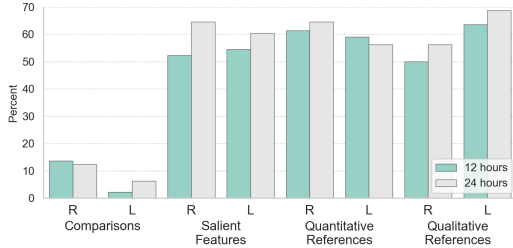

Fig. 5: Coded observations in percentage of all text responses in task 1 for radial (R) and linear (L).

In Fig. 6, we can furthermore observe some differences between the four conditions with respect to the time periods mentioned in the users' observations. First, it is clearly visible that the rush hour peaks were most commonly reported with the 24-hours variants, in particular 24l. Second, using 12l, users were reporting fewest observations about the afternoon rush hour. Third, using 12r, users reported few observations in the early morning, but more in the late morning than with any other visualization. Users did not report more about late evening or night hours using 24r, despite its continuous mapping across midnight.
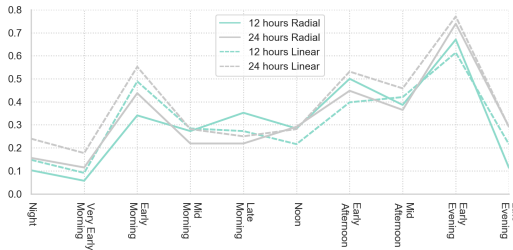

Fig. 6: Mean number of observations mentioning time periods of the day for the four conditions.

Hypothesis **H1** therefore is not supported: *The 12-hours radial chart imitating a clock layout does not lead to a larger number of insights when viewing daily patterns for a few seconds. On the contrary, users reported slightly more observations and also used qualitative references to points in time (e.g., "evening") significantly more often with a linear layout than with a radial layout.*

5.2 Task 2: Locate Time

We first computed the error by comparing the user's selected bar to the actual bar corresponding to the displayed label. From the 184 total obtained samples 152 (82.6%) were correct. The highest error rate was detected for 12r (30%), and the lowest for 12l (9%). Across all conditions, most of the errors (57%) were caused by a 12 hours shift from the actual bar, *i.e.*, the users selected the corresponding a.m. or p.m. value or vice versa, while 23% of responses are shifted by a single hour, and 10% by a combination of switching a.m. and p.m. and shifting one hour (*i.e.*, 11 or 13 hours shift). Most of the single hour shifts arose using 24-hours variants (5 out of 7), and most a.m. and p.m. swaps occurred with 12r (8 out of 20). While these a.m./p.m. swaps by 12r were evenly distributed in both directions, all three a.m./p.m. swaps
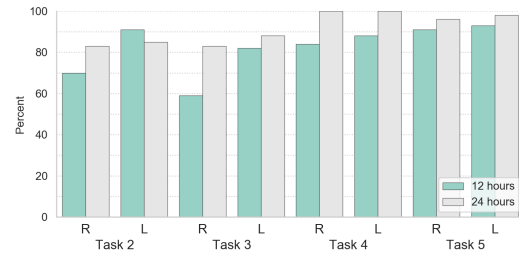

Fig. 7: Percentage of correct responses in tasks 2 — 5 for radial (R) and linear (L).

registered for 12l were incorrect selections from the a.m. chart, *i.e.*, the upper chart, while the requested time interval was given in p.m.

For the completion time of correct responses, we found a large and significant main effect for layout ($F_{1,61} = 21.279; p < .001; \eta_p^2 = .259$) but not for cardinality ($F_{1,61} = .223; p = .638; \eta_p^2 = .004$). On average, users needed 7.4 seconds to select the correct bar with radial charts and 5.2 seconds for linear charts. There is some trend for 24-hours charts to be more effective than two 12-hours charts when using a linear layout (15%, on average). There is a tendency for this trend to be reversed for radial charts, where the 24-hours variant was around 11% less effective in terms of completion time than the 12-hours variant. This interaction is statistically significant with a small effect size ($F_{1,61} = 4.504; p = .038; \eta_p^2 = .069$). This means that our hypothesis **H2** is only partially supported: *Juxtaposed 12-hours radial charts are slightly more efficient to locate a given time than a 24-hours radial chart, but they are more error-prone and less efficient than linear charts.*

5.3 Task 3: Read Value

From the total 184 samples, 50 input values corresponded to the actual data value. However, only 11 error cases differed by a value more than 4 from the actual data value, which is more than 10% of our maximum data values. We therefore counted all input values as error if they had an absolute deviation from the actual value of more than 4 (11 samples) or were invalid (*i.e.*, empty or non-numeric, 29 samples). The lowest error was measured for 24l (14%) and the highest for 12r (41%!). From the 18 error cases of 12r, 12 were invalid inputs. When analyzing the valid error cases across all conditions, we could explain 45% by rounding errors, where the user's input value lies within the same grid lines as the target value. The remaining error cases (four for 12r and one for 24r) can be potentially explained by one of the two following observations (Fig. 9): One error cause could be a misinterpretation of the visual encoding so that users reported the value of the closest bar to the axis label instead of the marked bar. A second error cause may be the misinterpretation of grid lines, where a wrong assumption of the grid line intervals can cause incorrect value interpretations.

To assess the efficiency of the charts, we performed an ANOVA on completion time of the 64 users not committing any errors. We found a large main effect for layout ($F_{1,62} = 21.064; p < .001; \eta_p^2 = .254$), but not for cardinality ($F_{1,62} = .220; p = .641; \eta_p^2 = .004$) and no interaction between the two factors ($F_{1,62} = .404; p = .527; \eta_p^2 = .006$). Overall, 12r required most time to come up with the correct response, and 24l led to the fastest response. This partially supports hypothesis **H3**: *Reading values from a linear chart is less error-prone than reading values from a radial chart and also more efficient. However, there is no effective difference between 12-hours and 24-hours linear charts.*

5.4 Task 4: Locate Maximum

From all 184 cases, two responses were invalid and excluded from further analysis. From the remaining 182 cases, there were 170 correct responses. For 24l and 24r, all recorded responses were correct. For 12l, 88% of the responses were correct, for 12r 84%. We qualitatively analyzed all 12 error cases and classified the potential reasons for the errors into three categories. The most common reason (five for 12l
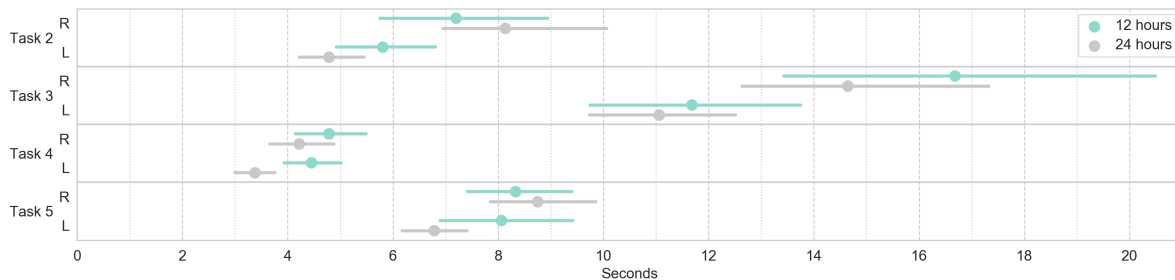
Fig. 8: Mean completion times with 95% confidence intervals in task 2 (locate time), task 3 (read value), task 4 (locate maximum), and task 5 (compare a.m./p.m. interval) for radial (R) and linear (L).
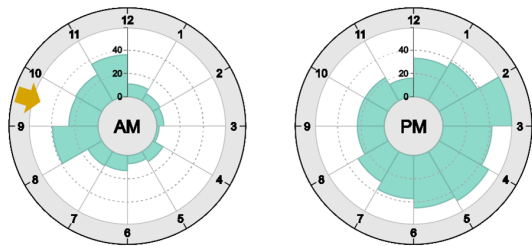

Fig. 9: Example for incorrectly read value with 12r (correct value is 25, user input was 7), which could be caused by an assumed grid line interval of five or by reading the value of the bar at 12-1 a.m.

and four for 12r) was that users selected the maximum value from the incorrect chart – *i.e.*, from a.m. instead of p.m. For the three remaining error cases of 12r, users selected either the second highest bar (two cases) or the highest value next to one of the value axes (one case).

The ANOVA on completion time showed small but significant main effects for layout ($F_{1,79} = 7.507; p = .008; \eta_p^2 = .087$) and for cardinality ($F_{1,79} = 5.721; p = .019; \eta_p^2 = .068$), but no interaction ($F_{1,79} = 2.742; p = .102; \eta_p^2 = .034$). On average, the correct maximum was selected with a linear layout in 3.9 seconds and with a radial layout in 4.5 seconds. Cardinality 24 led to significantly faster selection times (3.8 seconds) than cardinality 12 (4.6 seconds). Overall, the maximum was selected fastest using 24l (3.4 seconds). Hypothesis **H4** is therefore supported: *Identifying the maximum value is most efficient using a single, linear 24-hours representation of the time series data.*

5.5 Task 5: Compare AM/PM Interval Values

We counted every incorrect response as error, which are 10 cases in total. We found the highest number of incorrect responses for 12r (9%) and the lowest for 24l (2%). In four cases, users reported that the indicated values are equal when their value differences were rather low (1-7). This means, users underestimated the value differences. These error cases only happened with 12r (2) and 12l (2).

The ANOVA on completion time showed a medium-sized main effect for layout ($F_{1,80} = 11.516; p = .001; \eta_p^2 = .126$), but not for cardinality ($F_{1,80} = .216; p = .644; \eta_p^2 = .003$). We also found a medium-sized interaction between the two factors ($F_{1,80} = 8.393; p = .005; \eta_p^2 = .095$). On average, the correct comparison was performed within 7.3 seconds for linear layouts and 8.6 seconds for radial layouts. The fastest response was given with 24l (6.8 seconds), the slowest with 24r (8.8 seconds). The completion time of the two 12-hours variants was very similar (8.1 and 8.3 seconds). Therefore, hypothesis **H5** is partially supported: *Linear 24-hours charts facilitate the comparison of a.m./p.m. interval values. The 12-hours linear chart is not superior to the radial charts.*

5.6 Task 6: Subjective Rating

For the ratings, we found a large main effect for layout ($F_{1,88} = 83.912; p < .001; \eta_p^2 = .488$), but not for cardinality ($F_{1,88} = .548; p = .461; \eta_p^2 = .006$), and no interaction ($F_{1,88} = .156; p = .694; \eta_p^2 = .002$). On average, the linear charts received ratings 1.2 points higher than the radial charts on the 5-point Likert scale (see Fig. 10). 24l received the highest average rating (4.3) and 12r lowest (2.95).
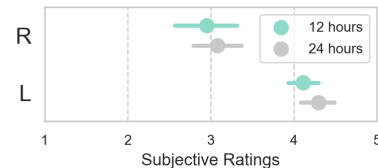

Fig. 10: Mean ratings issued on a five-point Likert scale with 95% confidence intervals.

To better understand the ratings, we analyzed the users' optional text comments. We received 66 text comments in total – most for 24r (21, which is a 43% comment rate) and fewest for 12l (27% comment rate). 24r received most purely negative comments (66%), followed by 12r (50%). The 24-hours radial chart was mainly criticized for its unusual appearance with comments like: *"It's actually a useful way to show the info, but is unusual and could be confusing at first."* or: *"It is not easy to interpret. I know it follows a clock format but it doesn't translate well."* The 12-hours radial charts received similar comments, like: *"This is really hard to look at until you realize that it is showing like a clock. But the data is still difficult to decipher."*

12l received more purely positive comments (50%). 24l did not receive any purely negative comments. Compared to the juxtaposed 12-hours radial charts, the linear bar charts were easier to read, e.g.: *"Now that I've seen both, this method is easier and more intuitive than the clock type visualization. It's easier to look at the bar charts and digest the information."* The 24-hours bar chart was primarily appreciated for its effectiveness and simplicity, for instance because *"it seems clear and simple, very easy to comprehend at a glance."* A list of all user comments can be found in the supplemental material. This only partially confirms hypothesis **H6**: *Both radial charts (12r and 24r) received significantly lower subjective ratings and more negative comments than the linear charts.*

6 DISCUSSION

Our main hypothesis that daily temporal patterns can be interpreted more effectively using clock-like 12-hours rose charts is not supported. Overall, the conventional and most elementary continuous linear bar chart performed more accurately and efficiently across all low-level tasks. This is particularly remarkable as the linear 24-hours bar chart requires least ink and space to encode the data. We discuss the results with respect to the experimental factors and our main hypothesis.

6.1 Linear vs. Radial Layout

Summarized, the independent variable layout can explain most observed performance differences:

- Using linear charts, users reported more observations about the data, and these observations more often used qualitative time references than observations derived from radial data visualizations.

- All low-level tasks could be solved significantly faster using a linear layout than both radial layouts, including finding a specified hour of a day.

- Participants clearly preferred linear charts. Many users reported that radial charts are harder to read, and some even reported fundamental difficulties interpreting the visual encoding.

To a certain extent, these results were expected and generalize findings that reading values, locating the maximum and comparing values is less accurate and efficient using radial bar and line charts [11,29] to rose charts. After analyzing error cases and user feedback, we speculate that rose charts are sometimes decoded fundamentally wrong, for instance by not knowing how the indicated bars' length corresponds to the axis on the top or by assuming a wrong grid line interval. For instance, one user reported that *"I JUST noticed that each bar represents 20 accidents. Before I thought 1 bar = 1 accident."* This can also explain the slightly lower performance in task 1, as some users probably had problems interpreting the visualized data without guidance how to read the chart.

Radial charts are not uncommon. Yet, it seems that conventional linear bar charts are much more ubiquitous than radial charts – even more than classic pie charts [48]. Consequently, they probably can be understood by a broad audience without any instructions. This is supported by various users' feedback, such as *"this is a form that's more familiar – no time really needed to understand the format of the data"* [24l].

On the other hand, some users also commented that they found radial charts more appealing or interesting with comments like *"this* [24l] *is simpler, less pretty or fancy but more effective than the previous one"* [24r] or *"[...] I just think the other one* [24r] *is much more fun."* This supports Borkin et al. [8] who speculate that traditional linear charts are less engaging, which could explain why they are least memorable for a non-expert audience. For hierarchical data, Merčun [41] also showed that linear, intended lists are perceived as more useful and efficient, but less engaging and appealing than radial sunburst and circle pack diagrams. These findings have a high relevance for the field of infographics, where designers seek to create aesthetically pleasing visual representations and less often use conventional bar or line charts [8]. Our results reinforce that designers of visualizations for the masses should stick to very simple and ubiquitous visual encodings to lower the risk of data misinterpretation due to incorrect decoding.

6.2 2×12-Hours vs. 1×24-Hours Chart

The separation into two less cluttered charts had much less influence on the completion time results than the layout. However, it has some negative influence on the data interpretation:

- Locating the maximum is most efficient using a continuous, linear chart and also more error-prone using separate charts. This is also indicated by a slightly higher number of salient features reported using 24-hours variants in the untargeted analysis.

- There is a small tendency to underestimate value differences between bars in two separate charts.

Even though the linear 24-hours bar chart was the only visualization that did not benefit from the fact that we compared associated a.m./p.m. intervals, it performed better than the radial charts and the separate 12-hours linear bar charts in task 5. For both, linear and radial separate charts, we found that users underestimated value differences. This

reinforces that lengths comparisons are easier when the bars are horizontally aligned, thereby confirming previous results [15,31,47,50].

It seems that users directed more attention to the a.m. data using the linear 12-hours chart than with any other visualization. This is indicated by the following observations with 12l: 1) users reported relatively more observations in the morning than in the afternoon in task 1 (Fig. 6), 2) some users mistakenly selected a.m. intervals instead of p.m. intervals in task 2, but never vice versa, and 3) users sometimes selected the maximum value of the a.m. chart instead of the p.m. chart in task 4, but never vice versa. Diehl et al. [18] report a reading direction effect from top-left to bottom-right in Cartesian charts, which could be an explanation for these findings. We therefore assume that the performance difference between the 12- and 24-hours variants of linear charts was caused by the separation of a.m. and p.m. values rather than the different bar widths. Separating a bar chart into meaningful sub-units therefore can be a way to direct the user's attention to the data encoded in the top chart. However, if accurate value comparisons or understanding of global maxima are required, a continuous, linear chart is the best option.

6.3 Clock Layout

Finally, we discuss the results concerning the hypothesis that clock-like charts facilitate the cognition of daily patterns:

- Across all low-level tasks, users committed most errors using the two juxtaposed 12-hours radial charts.

- Locating one-hour day-time intervals is slightly more efficient using two 12-hours radial charts than one 24-hours radial chart, but more error-prone and less efficient than if using linear charts.

- The 24-hours radial chart received more negative subjective feedback than the 12-hours radial chart.

One particularly unexpected finding was that users were significantly more efficient locating a given one-hour time interval using a linear layout than using a radial layout. This contradicts results by Fuchs et al. [26], which indicate that a radial layout is more suitable for detecting day-time locations. Partially, the poor performance of radial charts can be explained by the high completion times for the 24-hours radial chart (Fig. 8, first row). This was expected, since the 24-hours radial chart does not use the common clock layout. A major difference to the study by Fuchs et al. [26] is that our users were presented with large-scale visualizations instead of a matrix of glyphs. Therefore, our visualizations had labels indicating the precise hour of the day, which can be an explanation why users could locate the specified time efficiently in the linear charts compared to the unlabeled linear stripe and line charts by Fuchs et al.

Most of the errors when locating the time using the radial 12-hours variant was caused by swapping a.m. and p.m. This error also occurred with the linear 12-hours chart, but less frequently, and only if the time interval to be selected was p.m. In contrast to the two linear bar charts, two juxtaposed rose charts do not support the natural reading order. In the original rose charts of the mortality in the Crimean War, Florence Nightingale connected March and April 1855 in her two juxtaposed charts with a line. In our study, the two 12-hours rose charts were implicitly connected by noon *and* midnight, which are located on the top. This ambiguous meaning of 12 on the top was perceived as difficult to understand: *"I found it very hard to read initially as I was trying to figure out if the 12pm on the left was for PM or AM, based on it stating AM in the center. This is very confusing."* We also observed an increased tendency to report time periods just before noon in task 1 using the separate 12-hours charts compared to the other visualizations (Fig. 6). This could indicate that the users' attention is strongly focused to the point in time where the chart is separated – even though there is no salient data point. We speculate that this effect stems from the artificial discontinuity between the non-neighboring data points rendered next to each other (e.g., Fig. 1a). To systematically test this hypothesis, an eye tracking study comparing separate rose charts with changing variability between neighboring bins will be an interesting future work.

The separation of the 24 data bins into two juxtaposed charts also affects the appearance of the bars. The segments of the single 24-hours variant were narrower and the blue areas therefore looked more like bars than circle segments. Maybe it was perceived – and therefore also read – like a traditional bar chart by more users.

From the user feedback, we suspect that some users also did not draw the connection between a clock layout and the 12-hours rose charts – at least not initially. But even users who explicitly stated that they did see the connection between a clock and the rose chart reported that they found it hard to read. Rose charts for daily patterns use the same layout but a different encoding than analog clocks. It could be that people are trained to read the orientation of the clock hand rather than to memorize the circle segment corresponding to a particular hour. This way, they cannot translate their clock reading skills to interpret daily patterns or even to locate a given hour on a ring separated in twelve segments. Reading the clock itself is also considered a cognitively demanding task, and a common task to assess cognitive impairments is to ask patients to draw clocks showing a given time [43].

Only few users explicitly appreciated the clock layout, for instance *"I like this visualization because I can easily understand it based on intuition. It looks like a clock."* or *"I think reading particular numbers of this graph* [12l] *is easier, but the clock graph* [12r] *gives a better feel for time of day."* However, some users also associated the 24-hours radial chart with a clock: *"It's neat in that it parallels an analog clock face."* It therefore seems that it is mainly the radial layout that creates the impression of reading a clock, not necessarily the radial arrangement of 12 hours.

In summary, 12-hours radial charts combine the disadvantages of radial charts (inaccurate value reading, inefficient location of global maxima, inefficient value comparisons, and unfamiliarity) and separate charts (higher risk of reading the wrong chart and tendency to underestimate value differences). In addition, the meaning of the top 12 is ambiguous, and the bars get more distorted due to the larger circle segments compared to the 24-hours variant. The connection to the clock layout was also not apparent to many users and even if so, they reported problems utilizing the clock reading skill to interpret rose charts. The assumed intuitiveness of visualizing daily patterns in an analog clock layout could not be shown in this study.

7 STUDY LIMITATIONS

A validity threat of every crowd-sourced study is the lack of control over the **environment**. One potential confounding factor in our study is the browser window size of the users, which affects the four conditions differently, depending on the size of the visualization. We did not query the window size, but even the largest visualization (*i.e.*, 12r) is fully visible on a common mobile phone screen in portrait mode. Also, we have very few error cases in task 5, where both charts had to be interpreted. We can assume that both charts were visible for our users.

We selected one radial **encoding** to visualize daily patterns which is inspired by very early work by Florence Nightingale [10, 16]. However, there are other encodings that can be used with a linear and radial layout (Table 1). Using color, for instance, there is also the possibility to use a spiral display [19, 51], which is also interesting as it manages to overcome the problem of two juxtaposed charts.

As our goal was to simulate charts as found in news or public websites, we used a fixed, representative **size** for all visualizations. However, temporal patterns are sometimes also presented as glyphs [26] or sparklines [5]. As these small-scale visualizations usually only convey a general trend, the low-level analysis tasks used in our experiment are not appropriate for these visualizations.

We used only one type of **data** in our study. Based on prior results showing that different data sets do not interact with the layout [11], we assume that data with the same characteristics (e.g., daily network traffic [22, 37], mobile phone usage [46], or traffic information [34]) would lead to comparable results. However, previous work suggests an interaction between layout and bin size for different time periods [11, 26], where increasing the data granularity and the number of bins, respectively, improves the relative performance with radial charts. Thus,

it would be interesting to investigate daily patterns showing a finer granularity, such as 15-minutes bins, for radial charts in the future.

Finally, we performed our experiment with non-expert **users**. Subjective feedback indicates that participants got more used to radial charts over the short duration of the experiment. However, the self-reported visualization literacy of our participants did not have an influence on the number of reported observations or subjective ratings. It is therefore unclear whether rose charts are suitable for visualizing daily patterns for trained expert users. Also, some users may not be confronted with analog clocks so frequently and may require more training to get familiar with the clock layout.

8 CONCLUSIONS

Complementing prior work on radial charts for visualizing monthly and weekly periodical data [11, 26], we showed that radial charts also do not improve the understandability of *daily* periodical data – despite their resemblance with an analog clock. Our study also extends previous results, which showed that radially arranged bars can be decoded less accurately and efficiently than linearly arranged bars in many low-level analytical tasks [11, 29], to rose charts. Furthermore, we showed that chart separation makes the detection of salient features slightly more difficult, and may lead to underestimation of value differences. Based on these observations, we provide two recommendations for designers.

First, visualizations for the masses should rely on linear bar charts instead of radial rose charts even for showing periodical daily data. Shaping a rose chart like a clock does not improve the users' understanding of daily patterns.

Second, continuous bar charts should be preferred over separate ones, especially when accurate data interpretation is important. However, when horizontal space is limited, it is more advisable to separate a linear bar chart into multiple charts rather than to use a single rose chart. Separation of radial charts should be avoided, as it can cause unwanted attention effects.

In summary, we have contributed further evidence against the usage of radial charts, in particular rose charts. However, there might be other data and application domain cases, such as the analysis of wind directions, where rose charts can be highly beneficial.

REFERENCES

[1] M. Adnan, M. Just, and L. Baillie. Investigating time series visualisations to improve the user experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 5444–5455. ACM, 2016.

[2] D. Albers, M. Correll, and M. Gleicher. Task-driven evaluation of aggregation in time series visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 551–560. ACM, 2014.

[3] B. Alsallakh, W. Aigner, S. Miksch, and H. Hauser. Radial sets: Interactive visual analysis of large overlapping sets. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2496–2505, 2013.

[4] K. Bale, P. Chapman, N. Barraclough, J. Purdy, N. Aydin, and P. Dark. Kaleidomaps: a new technique for the visualization of multivariate time-series data. *Information Visualization*, 6(2):155–167, 2007.

[5] F. Beck and D. Weiskopf. Word-sized graphics for scientific texts. *IEEE Transactions on Visualization and Computer Graphics*, 23(6):1576–1587, 2017.

[6] T. Blascheck, L. Besançon, A. Bezerianos, B. Lee, and P. Isenberg. Glanceable visualization: Studies of data comparison performance on smartwatches. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):630–640, 2018.

[7] R. Borgo, L. Micallef, B. Bach, F. McGee, and B. Lee. Information visualization evaluation using crowdsourcing. In *Computer Graphics Forum*, vol. 37, pp. 573–595. Wiley Online Library, 2018.

[8] M. A. Borkin, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, and H. Pfister. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2306–2315, 2013.

[9] G. Boulton-Lewis, L. Wilss, and S. Mutch. Analysis of primary school childrens abilities and strategies for reading and recording time from analogue and digital clocks. *Mathematics education research journal*, 9(2):136–151, 1997.

[10] L. Brasseur. Florence nightingale's visual rhetoric in the rose diagrams. *Technical Communication Quarterly*, 14(2):161–182, 2005.

[11] M. Brehmer, B. Lee, P. Isenberg, and E. K. Choe. Visualizing ranges over time on mobile phones: a task-based crowdsourced evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):619–629, 2019.

[12] M. Burch and D. Weiskopf. On the benefits and drawbacks of radial diagrams. In *Handbook of Human Centric Visualization*, pp. 429–451. Springer, 2014.

[13] E. Burny, M. Valcke, A. Desoete, and J. E. H. Van Luit. Curriculum sequencing and the acquisition of clock-reading skills among chinese and flemish children. *International Journal of Science and Mathematics Education*, 11(3):761–785, 2013.

[14] M. C. Chuah. Dynamic aggregation with circular visual designs. In *Proceedings of IEEE Symposium on Information Visualization*, pp. 35–43. IEEE, 1998.

[15] W. S. Cleveland and R. McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American statistical association*, 79(387):531–554, 1984.

[16] I. B. Cohen. Florence nightingale. *Scientific American*, 250(3):128–137, 1984.

[17] Common Core State Standards Initiative. Mathematics Standards – Grade 1 – Measurement & Data – Tell and write time. http://www.corestandards.org/Math/Content/1/MD/B/3/, 2019. [Online; accessed June-2019].

[18] S. Diehl, F. Beck, and M. Burch. Uncovering strengths and weaknesses of radial visualizations—an empirical approach. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):935–942, 2010.

[19] P. Dragicevic and S. Huot. Spiraclock: a continuous and non-intrusive display for upcoming events. In *CHI'02 Extended Abstracts on Human Factors in Computing Systems*, pp. 604–605. ACM, 2002.

[20] G. M. Draper, Y. Livnat, and R. F. Riesenfeld. A survey of radial methods for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 15(5):759–776, 2009.

[21] Y. Drocourt, R. Borgo, K. Scharrer, T. Murray, S. Bevan, and M. Chen. Temporal visualization of boundary-based geo-information using radial projection. In *Computer Graphics Forum*, vol. 30, pp. 981–990. Wiley Online Library, 2011.

[22] R. F. Erbacher, K. L. Walker, and D. A. Frincke. Intrusion and misuse detection in large-scale systems. *IEEE Computer Graphics and Applications*, 22(1):38–47, 2002.

[23] S. Few and P. Edge. Save the pies for dessert. *Visual Business Intelligence Newsletter*, pp. 1–14, 2007.

[24] F. Fischer, J. Fuchs, and F. Mansmann. Clockmap: Enhancing circular treemaps with temporal glyphs for time-series data. In *Proceedings of EuroVis Short Papers*, pp. 97–101. Eurographics, 2012.

[25] W. J. Friedman and F. Laycock. Children's analog and digital clock knowledge. *Child Development*, pp. 357–371, 1989.

[26] J. Fuchs, F. Fischer, F. Mansmann, E. Bertini, and P. Isenberg. Evaluation of alternative glyph designs for time series data in a small multiple setting. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3237–3246. ACM, 2013.

[27] J. Fuchs, P. Isenberg, A. Bezerianos, F. Fischer, and E. Bertini. The influence of contour on similarity perception of star glyphs. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2251–2260, 2014.

[28] A. Gogolou, T. Tsandilas, T. Palpanas, and A. Bezerianos. Comparing similarity perception in time series visualizations. *IEEE transactions on visualization and computer graphics*, 25(1):523–533, 2019.

[29] J. Goldberg and J. Helfman. Eye tracking for visualization evaluation: Reading values on linear versus radial graphs. *Information visualization*, 10(3):182–195, 2011.

[30] R. L. Harris. *Information graphics: A comprehensive illustrated reference*. Oxford University Press, 2000.

[31] J. Heer and M. Bostock. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 203–212. ACM, 2010.

[32] J. Heer, M. Bostock, V. Ogievetsky, et al. A tour through the visualization zoo. *Commun. Acm*, 53(6):59–67, 2010.

[33] K. Hinum, S. Miksch, W. Aigner, S. Ohmann, C. Popow, M. Pohl, and M. Rester. Gravi++: Interactive information visualization to explore highly structured temporal data. *Journal of Universal Computer Science*, 11(11):1792–1805, 2005.

[34] X. Huang, Y. Zhao, C. Ma, J. Yang, X. Ye, and C. Zhang. Trajgraph: A graph-based visual analytics approach to studying urban network centralities using taxi trajectory data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):160–169, 2016.

[35] INFOweTRUST. Creative Routines. https://infowetrust.com/creative-routines/, 2014. [Online; accessed Dec-2018].

[36] W. Javed, B. McDonnel, and N. Elmqvist. Graphical perception of multiple time series. *IEEE Transactions on Visualization and Computer Graphics*, (6):927–934, 2010.

[37] C. Kintzel, J. Fuchs, and F. Mansmann. Monitoring large ip spaces with clockview. In *Proceedings of the 8th International Symposium on Visualization for Cyber Security*. ACM, 2011.

[38] R. Kosara and C. Ziemkiewicz. Do mechanical turks dream of square pie charts? In *Proceedings of the 3rd BELIV'10 Workshop: Beyond Time and Errors: Novel Evaluation Methods for Information Visualization*, pp. 63–70. ACM, 2010.

[39] J. Lankow, J. Ritchie, and R. Crooks. *Infographics: The power of visual storytelling*. John Wiley & Sons, 2012.

[40] M. Masoodian, B. Endrass, R. Buhling, P. Ermolin, and E. André. Time-pie visualization: Providing contextual information for energy consumption data. In *Information Visualisation*, pp. 102–107. IEEE, 2013.

[41] T. Merčun. Evaluation of information visualization techniques: analysing user experience with reaction cards. In *Proceedings of the 5th BELIV'14 Workshop: Beyond Time and Errors: Novel Evaluation Methods for Visualization*, pp. 103–109. ACM, 2014.

[42] D. Ren, B. Lee, and M. Brehmer. Charticulator: Interactive construction of bespoke chart layouts. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):789–799, 2019.

[43] D. R. Royall, J. A. Cordes, and M. Polk. Clox: an executive clock drawing task. *Journal of Neurology, Neurosurgery & Psychiatry*, 64(5):588–594, 1998.

[44] D. Seebacher, M. Miller, T. Polk, J. Fuchs, and D. A. Keim. Visual Analytics of Volunteered Geographic Information: Detection and Investigation of Urban Heat Islands. In *Symposium on Visualization in Data Science (VDS) at IEEE VIS 2018*, 2018.

[45] J. Sheidin, J. Lanir, T. Kuflik, and P. Bak. Visualizing spatial-temporal evaluation of news stories. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion*, pp. 65–68. ACM, 2017.

[46] Z. Shen and K.-L. Ma. Mobivis: A visualization system for exploring mobile data. In *2008 IEEE Pacific Visualization Symposium*, pp. 175–182. IEEE, 2008.

[47] D. Simkin and R. Hastie. An information-processing analysis of graph perception. *Journal of the American Statistical Association*, 82(398):454–465, 1987.

[48] I. Spence. No humble pie: The origins and usage of a statistical chart. *Journal of Educational and Behavioral Statistics*, 30(4):353–368, 2005.

[49] J. Stasko and E. Zhang. Focus+ context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations. In *Proceedings of IEEE Symposium on Information Visualization*, pp. 57–65. IEEE, 2000.

[50] J. Talbot, V. Setlur, and A. Anand. Four experiments on the perception of bar charts. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2152–2160, 2014.

[51] C. Tominski and H. Schumann. Enhanced interactive spiral display. In *SIGRAD 2008*, number 034, pp. 53–56. Linköping University Electronic Press, 2008.

[52] P. Troshenkov. NYC Hourly Car Accidents 2013-2016. https://www.kaggle.com/pavetr/nypdcollisions/version/1, 2017. [Online; accessed Dec-2018].

[53] M. Weber, M. Alexa, and W. Müller. Visualizing time-series on spirals. In *Infovis*, vol. 1, pp. 7–14, 2001.

[54] J. Zhao, F. Chevalier, and R. Balakrishnan. Kronominer: using multi-foci navigation for the visual exploration of time-series data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1737–1746. ACM, 2011.