symmetry

Article

A Perception Study for Unit Charts in the Context of Large-Magnitude Data Representation

Yun Lin [1,*], Yi Tang [1,*], Yanfei Zhu [2], Fangbin Song [1] and Wenzhe Tang [2]

1 School of Design Art and Media, Nanjing University of Science and Technology, Nanjing 210014, China
2 School of Mechanical Engineering, Southeast University, Nanjing 211189, China
* Correspondence: yunlin@njust.edu.cn (Y.L.); tangyi@njust.edu.cn (Y.T.)

Abstract: Unit charts are a common type of chart for visualizing scientific data. A unit chart is a chart used to communicate quantities of things by making the number of symbols on the chart proportional to the number of items represented. An accurate perception of the order of magnitude is essential to evaluating whether a unit chart can effectively convey information. Previous studies have primarily focused on perceptual properties at small order-of-magnitude scales or the efficacy of pictographs in unit charts. However, few researchers have explored the perceptual effectiveness of unit charts when representing large orders of magnitude. In this study, we performed a series of sampling measurements to investigate the visual–perceptual characteristics of unit charts when representing asymmetric interactions such as large-scale numbers. The results showed that under the restriction of the current conventional display medium, unit charts still offer a significant advantage over bar charts in a single-scale visual overview. However, this comes at the cost of a longer response time. Although this study constitutes basic research, accumulating evidence about how people reason about magnitudes beyond human perception is critical to the field of information science. This study may contribute to understanding how viewers perceive unit charts and the factors that influence graphical perception. This article provides some specific guidelines for designing unit charts that may be useful to visualization designers.

Keywords: unit charts; visualization evaluation; graphical perception; size and scale perception

check for updates

Citation: Lin, Y.; Tang, Y.; Zhu, Y.; Song, F.; Tang, W. A Perception Study for Unit Charts in the Context of Large-Magnitude Data Representation. *Symmetry* **2023**, *15*, 219. https://doi.org/10.3390/sym15010219

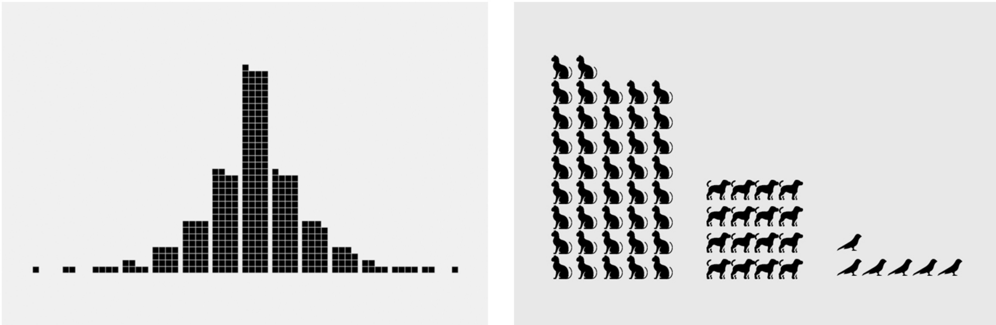
Academic Editor: José Ignacio Rojas Sola

Received: 18 November 2022
Revised: 6 January 2023
Accepted: 7 January 2023
Published: 12 January 2023
Corrected: 10 April 2023

1. Introduction

Representing values across multiple orders of magnitude at the overview stage is challenging. Taking a standard bar chart as an example, if the data are distributed between 0 and 10,000, viewers can easily perceive the ratio of 5000 and 7000, but it is challenging to perceive the ratio of 70 and 7000 or 60 and 90. Such graphical perception tasks are difficult due to the limitation of retinal resolution and the ability to take in and process orders of magnitude of differences. People have fundamental problems with orders of magnitude and calculations/interpretations on those scales. Decision-making across magnitudes may be different even without the errors due to visualization. It is beneficial to explore how data across large magnitudes can be effectively visualized to assist decision-making within the limitations of the display medium.

The motivation of this study is to explore the utility of unit charts when visualizing large magnitudes of data, thus providing a practical experience for visualization researchers and designers. Typically, charts represent values by variations of visual properties, such as bar charts by length, scatter charts by position, and bubble charts by area. Unit charts depict a value by multiples instead. The symbols in a unit chart usually encode a quantity, as shown in Figure 1. Compared with traditional charts such as bar charts and pie charts, unit charts can fully use display space and thus improve perceptual discriminability. Furthermore, due to the positive effects of pictorial visualization on memorability, engagement, and enjoyment [1–3], they have been increasingly incorporated into narrative visualization (e.g., infographics, data videos, data comics) to support data-driven storytelling for the general public [4]. However, few scholars have examined unit charts' estimation or counting issues due to the lack of widespread promotion.



a.a black chart b.a pictorial unit chart

Figure 1. Unit Charts.

In general, a well-designed unit chart requires both faithful data-binding and expressive visual design, which can be a challenging task even for experienced information designers. Previous perception studies related to unit charts have mainly focused on the impact of pictographs on information communication rather than magnitude estimation [2,5,6]. In this study, we will take a series of sample measurements to investigate the visual perception characteristics of unit charts based on previous research.

The remainder of this article is structured as follows. Section 2 reviews related research on unit charts, visualization of data across large magnitudes, and scale perception. Section 3 describes a psychophysical experiment measuring unit charts' just-noticeable difference (*JND*) and reports the experimental results. Section 4 reports on a user study that investigates the magnitude estimation for unit charts. Based on previous research conclusions, Section 5 analyses and discusses the results and limitations of the two experiments in this study. Section 6 describes the significance of our findings and provides recommendations for future work.

2. Related Works

2.1. Unit Charts

According to the definition by Robert L. Harris, a unit chart is a chart used to communicate quantities of things by making the number of symbols on the chart proportional to the number of items being represented [7]. Unit charts can be classified into two types based on the kind of symbols [7]. Occasionally, the chart is called a black chart when simple geometric or irregular shapes are used, see Figure 1a. The chart is usually referred to as a pictorial unit chart when the symbols are pictures, icons, or sketches, see Figure 1b.

This pictographic language can be traced back to the International System Of TYpographic Picture Education (ISOTYPE) proposed by Otto and Marie Neurath in the 1920s [8]. As a method of visual statistics, it consists of a set of standardized and abstracted pictorial symbols to represent scientific data [9]. The primary rule of ISOTYPE is that greater quantities are represented by a greater number of the same-sized pictogram rather than by an enlarged pictogram. This rule that combines the identical pictogram using serial repetition is also the most important in unit charts. The ISOTYPE has been widely accepted by the design community for its simple style and has been widely applied. For example, in recent years, pictograms/unit graphs have been used to display probability in medical decision-making, and these studies have demonstrated that this can help reduce cognitive bias [10,11]. In the past decade, researchers have studied pictorial visualizations to gain an understanding of how they are designed [3,12,13] and what benefits they may bring [1,2,14]. In terms of visual design, Boy et al. proposed a design space of anthropographics, which includes four dimensions: human shape, visualization, unit labeling, and unit grouping [12]. Morais et al. continued to further extend the design space of pictorial visualization by introducing the dimensions of granularity, coverage, coverage, realism, specificity, situatedness, and physicality [13]. Nana Wang and Leah Burns used ISOTYPE as a case study to examine the potential and limitations of pictorial diagrams for knowledge development and communication. They identified vital approaches and steps for creating ISOTYPE pictorial diagrams from the perspective of knowledge transfer and reasoning support [15].

Regarding the effectiveness of pictorial visualizations most relevant to this study, scholars still differed on this issue. Haroz et al. found that pictorial visualizations can improve working memory, engagement, and performance [2]. Hugo Romat and his colleagues investigated how pictographs influenced the visualization experience for personalization and immersion [3]. The results of crowdsourcing experiments conducted by Alyxander Burns et al. suggested that, at least for simple representations of parts-whole relationships, the choice of icon or symbol does not significantly impact meaning construction and insight extraction [14]. Amini et al. suggested that animation and pictographs can enhance the understandability of data insights and encourage viewer engagement in data-driven clips [1]. Günther Schreder et al. presented Isotype's positive effects of countability, iconicity, and ancillary semantic information on graph comprehension from a cognitive perspective [16]. On the contrary, Some scholars are very wary about using pictographs due to their distraction from the data itself [17]. It has been observed by Burns et al. that pictographs have little impact on understandability in part-to-whole visualizations [15]. Stephen Few criticized the inappropriate use of unit charts and argued that visualization designers should focus on communicating information in meaningful ways [18]. Stephen Few believed that even though segmentation may lead to more accuracy, it encourages people to slow down and count, which is a less efficient cognitive strategy. Trogu Pino questioned the rationality of this tedious strategy, which requires a viewer to extract information from a typical ISOTYPE chart by counting the symbols in each row and multiplying them by a given scale to obtain the total [19]. It is worth noting that Pino's criticisms relating to counting may not be how users actually engage with these charts, and thus this is not necessarily a fair characterization. It would seem that pictographs should reduce to area charts at large scale/distance, where the units blend into a single whole and area judgments are possible.

Through a systematic literature review of unit charts, it appears that most existing issues revolve around the utility of pictographic visualization when representing small-scale data scenarios; however, the magnitude perception characteristics of the unit charts remain largely unexplored. Although unit charts emerged about a hundred years ago, logical representation methods and knowledge transfer and reasoning support methods are still not obsolete. We should reflect on this in the context of the current information explosion.

2.2. Visualization of Data across Large Magnitudes

The multiscale visualization technique is popular for presenting data across large magnitudes. Multiscale visualizations allow users to present, navigate and relate data across multiple abstraction scales [20]. They are commonly used to analyze multiscale processes and data in various application areas, such as Radiation Dose Chart [21] which uses unit charts with different colors to show multiple different scales of radiation dose. Compared with single-scale visualization, multiscale visualization has the advantages of allowing more extensive data sets to be visualized, producing less confusion, and allowing patterns to emerge at different scales.

As users explore a dataset, they usually start with an overview, zoom in and filter, and then focus on details. In complex visual analysis, obtaining pattern and overview information at the same scale is one of the most fundamental tasks [22]. Logarithmic scales are commonly used to cover a wide range of values, but their non-linearity makes specific tasks difficult to comprehend [23]. Hlawatsch et al. presented a new approach for bar charts, which uses multiple scales to cover a large value range, while the linear mapping within each scale preserves the ability to visually compare quantitative ratios [24]. The scale-stack bar charts combine the advantages of linear and logarithmic scales while avoiding their drawbacks. Multivariate visual encoding, encoded with multiple visual variables, is another convenient way of improving graphical perception [25]. Moody mentioned that discriminability depends mainly on the visual distance between the symbols [26]. The visual distance is affected by the number of different visual variables and the size of their differences. However, the effect of different visual variables on graphical perception is not mutually independent. Garner proposed dimension integrity, which expounds that this effect is facilitated or interferential in different situations [27]. Overall, providing a comprehensive overview of a wider range of values in a limited display space remains challenging while ensuring good usability and user experience.

2.3. Size and Scale Perception

Estimating quantity is crucial for everyday life and success in the STEM disciplines [28]. Most of the research on people's reasoning about magnitude has concentrated on magnitudes at scales within human experience scales. People can instantly recognize the number of objects without counting them, an ability known as subitizing [29–31], up to a maximum of about 4–5 objects. As the number of items increases beyond that range, the visual system shifts to either slow counting or noisy estimation.

However, quantities outside human perception are difficult to reason about (e.g., nanoseconds and geologic time). Researchers are increasingly examining reasoning about larger scales [32–34]. There is a convergence of evidence from neurocognitive, cognitive, and developmental fields suggesting that reasoning about any type of magnitude (e.g., abstract, spatial, temporal) uses the same neural and conceptual resources (e.g., [35–39]). A general magnitude system [37] or a magnitude theory [38] contends that the inferior parietal cortex is responsible for processing all more/fewer judgments necessary for action. During processing, magnitude information is automatically extracted as a spatially organized series of numbers (e.g., [40–42]). An organized mental number line describes how magnitude is organized spatially (e.g., [42–44]). It is debated what distribution pattern magnitude follows along the mental number line [45,46]. However, it is widely accepted that people possess compressed representations of the unfamiliar and relatively larger magnitudes. The "compression" in numerical cognition is similar to a log scale mapping of numeracy, it leads to the common occurrence of overestimating relatively small magnitudes and underestimating relatively large magnitudes [47,48]. According to the category adjustment model, magnitude is stored as a hierarchical combination of metric and categorical information [49,50]. In the mental number line, more familiar and smaller magnitudes may constitute individual categories that occupy the bulk of the mental number line. In contrast, a wide range of unfamiliar and larger numbers may be encompassed by several categories, such as "big" and "really big," thus making it more difficult to differentiate between these larger numbers.

3. Experiment 1: Measurement for the Discrimination Threshold of Unit Charts

The just-noticeable difference (*JND*), also known as the difference threshold, is the minimum level of stimulation that a person can detect 50% of the time [51]. The *JND* is a statistical, rather than an exact quantity: from trial to trial, the difference that a given person notices will vary somewhat, and it is, therefore, necessary to conduct many trials to determine the threshold. This study used a different proportion, and reported the value of the "75% *JND*". Even though the standard eye has a minimum angle of view (AOV) of one arcminute [52], it has not been established to what extent viewers can detect differences in unit charts due to their complexity. The motivation of Experiment 1 is to measure the viewers' perceptual discrimination threshold of unit charts on a conventional display, along with a discussion of how unit charts facilitate perceptual discrimination.

3.1. Stimuli and Apparatus

Visual expressiveness was not a concern in this study, so we removed superfluous pictographs and labeled images in the symbolic stimuli as much as possible. As shown in Figure 2, a square with an "n" pixels length was encoded as a unit granularity. The unit squares were combined in stacks to represent the values of each order of magnitude. The value range of side length "n" was 3–11, considering the minimum requirement of unit block construction and the limitation of experimental screen resolution. Experiment 1 aimed to measure the unit chart's minimum length "n" that participants could discriminate when it represented different orders of magnitude.

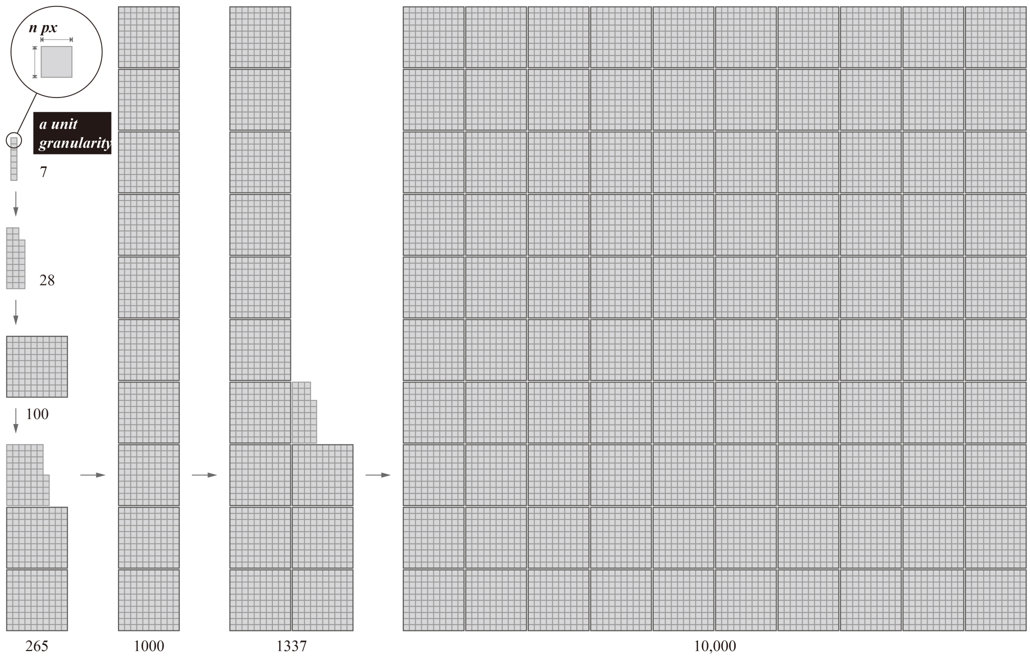


Figure 2. Symbolic Stimulus used in Experiment 1.

The visual stimuli and experimental program were generated using JavaScript. Experiment 1 was conducted on a computer running the Windows 11 operating system with a 4.90 GHz Intel Core i7 processor. The display was a 23.8-inch LCD monitor (Dell u2414 h) with a resolution of 1920×1080 pixels, a brightness of 250 cd/m^2, a static contrast of 1:1000, and a grayscale response time of 8 ms. We color-calibrated the LCD version of our display. The screen was placed approximately 70 cm from the participants' eyes and was made perpendicular to the participants' line of sight by adjusting the seat height. The experiment was conducted in a human factors laboratory under normal lighting conditions (about 300 lux).

3.2. Methodology

Experiment 1 used an estimate paradigm [53] to measure unit charts' *JND* through an adaptive psychophysical approach (a staircase procedure). Estimate paradigms required participants to estimate the value of continuous features in the stimulus directly. In the staircase procedure, we set it up to begin with a side length of six pixels (very detectable) and aim for the 75% threshold value. In this experiment, one step was defined as the edge length of the unit graph increasing or decreasing by one pixel. If the participant answered correctly, the next stimulus would decrease by one step while keeping the target value constant. The wrong choice would result in a one-step increase in the next stimulus, making the next estimation task more accessible, as shown in Figure 3. After each estimation, the target value of the following stimulus chart was randomly generated. Experiment 1 divided tested values into three intervals according to the orders of magnitude (tens, hundreds, and thousands). Ten randomly generated values in each interval were tested.

Through the staircase procedure, the *JND* was targeted by penalizing incorrect choices more than correct ones. These distance changes correspond to inferring "75%" *JND*s, or the minimum difference in value required to be reliably discriminated 75% of the time. The staircase procedure was terminated after reaching 50 individual estimations or meeting the convergence criteria. Based on the last 30 user estimations, the convergence condition was designed to determine whether discrimination between users was stable. Specifically, the 30 user estimations for each test chart were divided into three subgroups, and convergence was reached when there was no significant difference between these three subgroups by F-test (F (2, 27); $\alpha = 0.1$). In the final step, after the staircase procedure had been completed, we used the average distance in value between these subgroups as the *JND*.

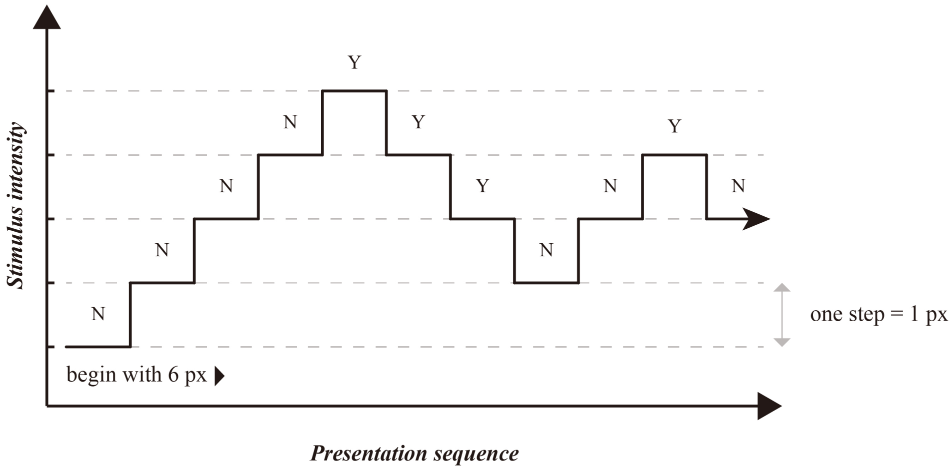


Figure 3. Staircase method. "Y" = Yes, the stimulus can be seen, and "N" = No, the stimulus cannot be seen.

3.3. Participants

A total of 30 participants (mean ± SD, 31.33 ± 3.29 years; 16 females, 14 males) of different ages and occupations were invited to participate in this experiment to avoid an overly homogeneous sample. Participants reported frequency of Internet use and proficiency in computer use, and over 65% of the participants could operate computers skillfully. All participants reported normal vision. A reward of approximately $15 was provided to each participant at the end of Experiment 1.

3.4. Procedure

Participants were required to complete the training and 15 estimation exercises before the start of the main trial. After each exercise trial, participants would receive feedback on whether their answers were correct. In formal trials, participants were required to fixate a black cross in the center of the screen for 500ms and then respond to the stimulus. Before the next loop started, a mask that would reduce afterimages was presented on the screen by 600 ms. The system checked whether the answer was correct or incorrect in each loop and appropriately assigned a response to end the trial. After completing a set of loops (by meeting convergence criteria), participants were given the option to take a short break and were informed of how many more trials they had left.

3.5. Results

The overview of collected *JND* data is shown in Figure 4a. The data were analyzed using SPSS statistical computer software (SPSS, Inc., Chicago, IL, USA). Statistical analysis included one-way ANOVA followed by Tukey-b or Tamhane tests, as required for each variable [54]. The box plot is shown in Figure 4b, and no outliers were identified. Post hoc multiple comparisons were determined using the Tamhane procedure, and the results showed that the mean *JND* of unit charts when representing tens digits (M = 7.35 px) is significantly lower than that when characterizing hundreds digits (M = 8.38 px) and thousands digits (M = 9.29 px). The results of multiple comparisons are shown in Table 1. Unit charts that represent larger magnitudes had larger mean *JND*s.

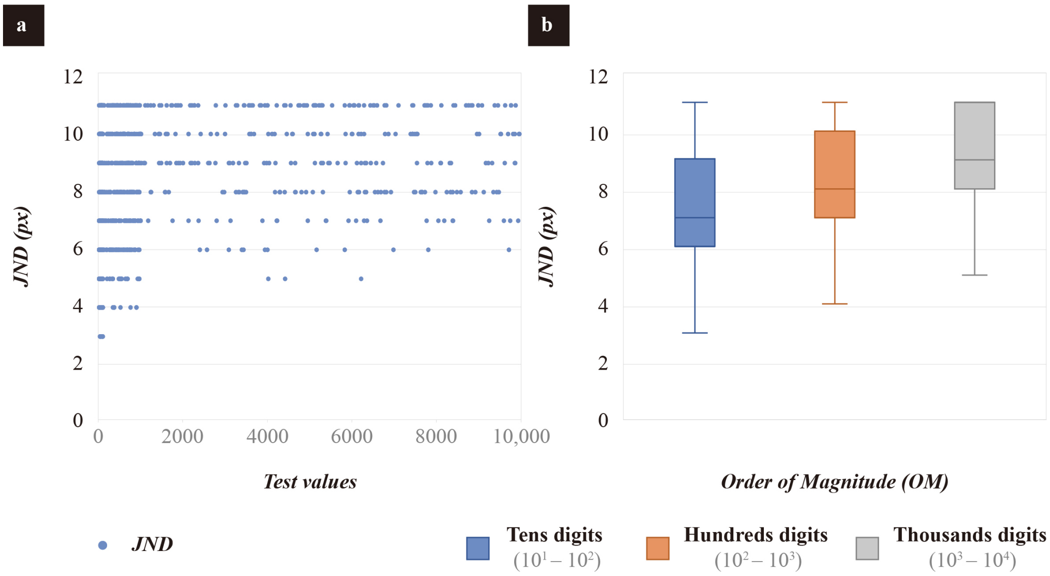


Figure 4. Measurement results of difference threshold. (**a**) Overview of the *JND* measurement results. (**b**) Box plot of *JND* for unit chart representing different orders of magnitude.

Table 1. Post hoc multiple comparisons with the Tamhane test for *JND*.

(I) OM	(J) OM	Mean Difference (I – J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Tens	Hundreds	−1.030 *	0.162	0.000	−1.420	−0.640
	Thousands	−1.940 *	0.148	0.000	−2.290	−1.590
Hundreds	Thousands	−0.910 *	0.143	0.000	−1.250	−0.0570

*. The mean difference is significant at the 0.05 level.

4. Experiment 2: A User Study of Magnitude Estimation for Unit Charts

The perceptual tasks in unit charts typically involve comparisons of magnitudes. Based on previous studies [2,15,55], this study focused on which perceptual pattern (counting, estimation, or mental arithmetic) users perform when using unit charts. To understand how they extracted information about the relationship between items from the different unit charts, Experiment 2 started with a small-scale web-based survey. In addition, we tested the effectiveness of two typical visualization tasks.

4.1. Stimuli and Apparatus

A number base in this study is the number of digits that a counting system uses to represent numbers. The base can be any whole number greater than zero. The most commonly used number system is the decimal system, commonly known as base 10. Its popularity as a counting system is most likely because we have ten fingers. To deeply explore the impact of the composition patterns on users' perceptual performance, we comprehensively examined two types of unit charts with base 10 and non–10. In this experiment, we chose 8 to represent the non–10 base, as shown in Figure 5. In the proportional estimation task, to examine whether order-of-magnitude differences affect task performance, we defined the difference in the order of magnitude between the larger (O^m) and smaller (O^n) values in the test chart set as the Diss_OM, whose value was the difference between the base power laws of the two values (m-n). We divided the ***Diss_OM (O^m, O^n)*** into three groups that differed by 1, 2, and 3 orders of magnitude. It should be noted that to ensure consistent graphical complexity, we divided the Diss_OM of the two test charts by corresponding bases. The equipment used in Experiment 2 is the same as that used in Experiment 1.

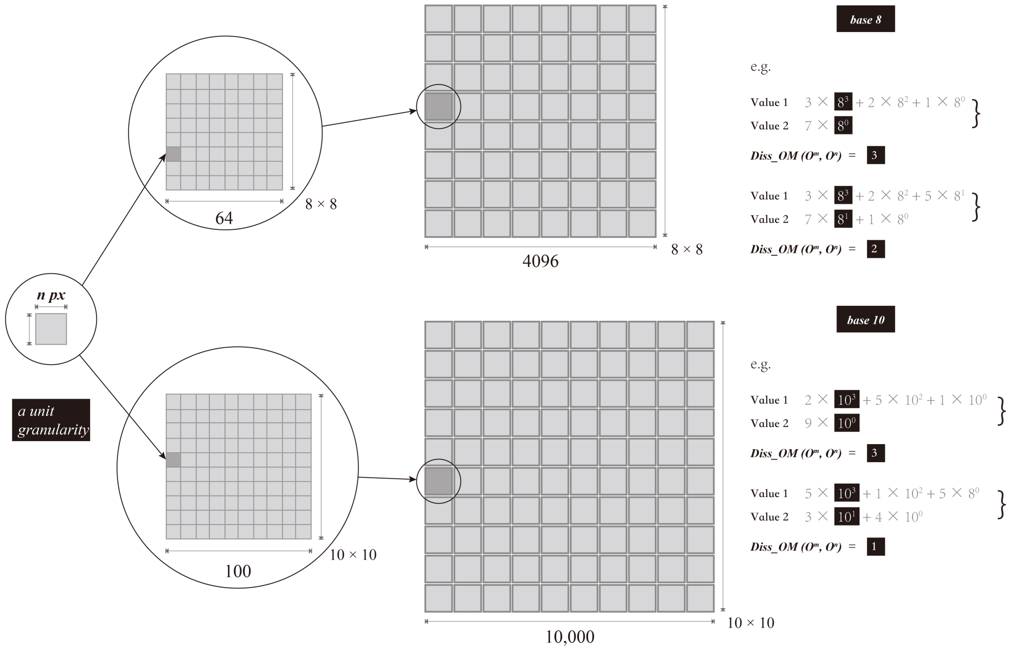


Figure 5. The base and the *Diss_OM* (O^m, O^n).

4.2. Participants

People have difficulty reasoning about magnitudes outside human perception [28], and participants' unfamiliarity and lack of knowledge of the relevant domain may confound the results. It was necessary for participants recruited for this experiment to have a visualization background and some expertise. We invited 30 undergraduates and postgraduates from a comprehensive research university majoring in visual communication design to participate in this experiment. All participants reported being proficient in using computers and using the Internet frequently. A total of 83% of the participants reported understanding visualization. Since this study focused on the perceptual mechanisms of viewers when using unit charts, choosing a homogeneous cohort was acceptable. The participants did not exhibit any color blindness or color weakness. Each participant received a reward of approximately USD 20 at the end of the experiment.

4.3. Perception Patterns Survey

We designed a questionnaire, as shown in Figure 6. A series of horizontal juxtaposed chart sets were presented in the questionnaire, and participants were asked to choose their preferred perception pattern (multiple choices were allowed) after viewing the test chart. Three perception patterns were visualized in the questionnaire, which are: (1) counting the precise values separately (Counting), (2) estimating the ratio of the area (Estimation), and (3) deriving the order-of-magnitude relationship by mental arithmetic (Mental Arithmetic). Participants were informed of the purpose and process before completing the survey questionnaire. All participants were informed in detail about the meaning of the three perception patterns. We also adopted a think-aloud protocol to record their thinking process as they complete the task. Participants are asked to say whatever comes into their minds including what they are looking at, thinking, doing, and feeling. The independent variables for the stimulus pairs we investigated included the two bases, and three Diss_OM mentioned above. Within each Diss_OM, three random combinations of values were presented, and each participant was required to view 2 × 3 × 3 = 18 stimuli.

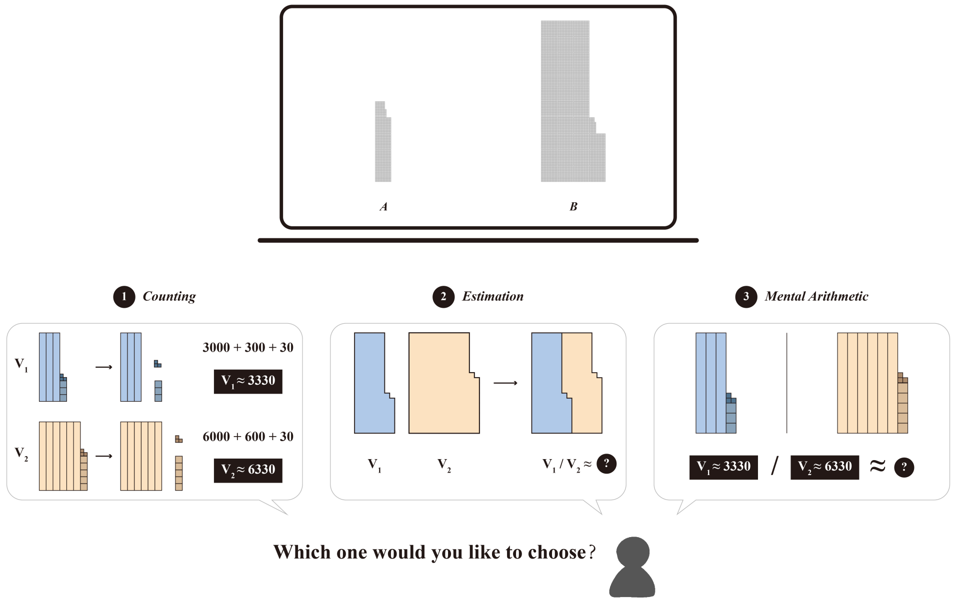


Figure 6. Diagram of the perception patterns survey.

A total of 30 questionnaires were collected. Figure 7 shows a summary of data, which can be summarized as follows: 1. For the base 8 unit charts, Estimation was the predominant perception pattern. We checked the stimuli corresponding to participants' choices of Counting and Mental Arithmetic, and they were both small values within 20, so relatively easy to be counted and calculated. 2. For the base 10 unit charts, many participants selected both Estimation and Mental Arithmetic. This is still open to discussion, as they may execute both Estimation and Mental Arithmetic during perception. Still, the key role may be only one of them, and it is difficult for participants to distinguish them. 3. We also observed that as Diss_OM became larger, more and more participants preferred to use Counting to directly report values. Some participants fed feedback that an overly large Diss_OM would make it more difficult for them to estimate. In addition, since V_1 is a single digit (very easy to count), their main effort was only to count the thousand-digit V_2, which were the two main reasons why they preferred Counting. We also found some interesting details from the transcribed think-aloud protocol. First, for the base 10 unit charts, participants would usually "optimize" the values in some way, as shown in Figure 8a, to achieve a more convenient estimate. Take 465/3337 as an example, participants would "optimize" to 460/3300, and furthermore, based on the multiplication table "4 × 8 = 32", they would roughly deduce that the ratio is 1/7. Second, participants partially used visual anchoring [56] in area comparison, as shown in Figure 8b. They would imagine a non-existent baseline in the center of the larger chart, and then "embed" the smaller chart for comparison.

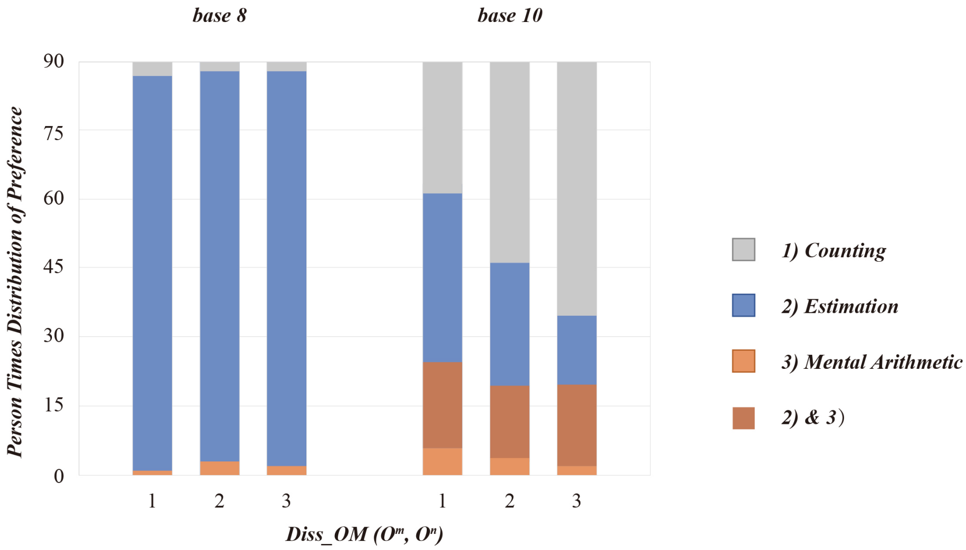


Figure 7. The summary of the types of perception patterns that users prefer to adopt when viewing unit charts.

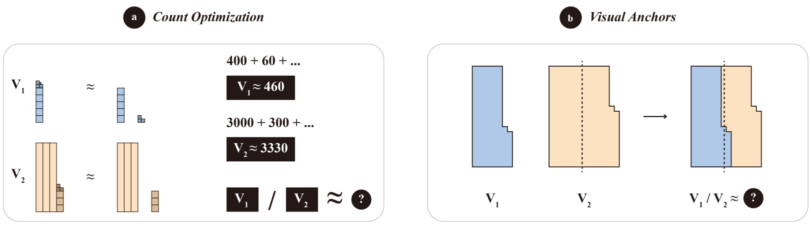


Figure 8. "Count Optimization" and "Visual Anchors".

4.4. Performance Measurements

Based on the survey results, we conducted a measurement experiment to quantify the performance of different visual perception tasks. Two visualization tasks for the Counting and Estimation pattern were developed. Task 1 was to count the values of the test charts as accurately as possible. Task 2 was to estimate the proportionality between the two test charts by area. It was clear from the user survey results that measuring the counting performance of the base 8 unit charts was too laborious to be meaningful, so only the base 10 unit charts were studied in this measurement experiment. For Task 2, the measurement experiment adopted a 2 × 3 within-subjects design, in which the variables were the base and the Diss_OM. Three values randomly generated by the experimental system within each difference grade interval were tested in Task 2.

The measurement experiment was divided into two groups according to the two task types. Before the formal experiment began, participants were required to complete a training session and three practice exercises. In Task 1, participants responded to the prompts in the interface by filling out the input boxes below the test chart as accurately as possible. Participants would receive feedback after completing each practice task. For Task 2, we provided two estimation forms for participants to choose from. Participants could directly enter the proportional values they estimated. Considering that mental arithmetic may be difficult for some participants, to obtain a more accurate mental representation of participants, we adopted a magnitude production method [53,57] that allowed participants to reflect their perceived proportions by adjusting an interactive juxtaposed barplot (see Figure 9). As the comparison in this experiment across multiple orders of magnitude, the bar chart may not be able to accurately adjust a certain proportion under a given situation, such as 2/3890, resulting in errors. Thus, we chose the juxtaposed barplot instead of a regular bar chart. Participants adjusted the juxtaposed barplot by moving the mouse, and the input box below would display the adjusted proportion value in real time. The stimulus chart coded with two variables appeared randomly in each trial. The experimental system recorded the correct answer, the participant's adjusted answer, and the response time during each trial. Participants were required to complete 81 estimation tasks and were given a short break after each set of trials. The measurements experiment took approximately 45 min.

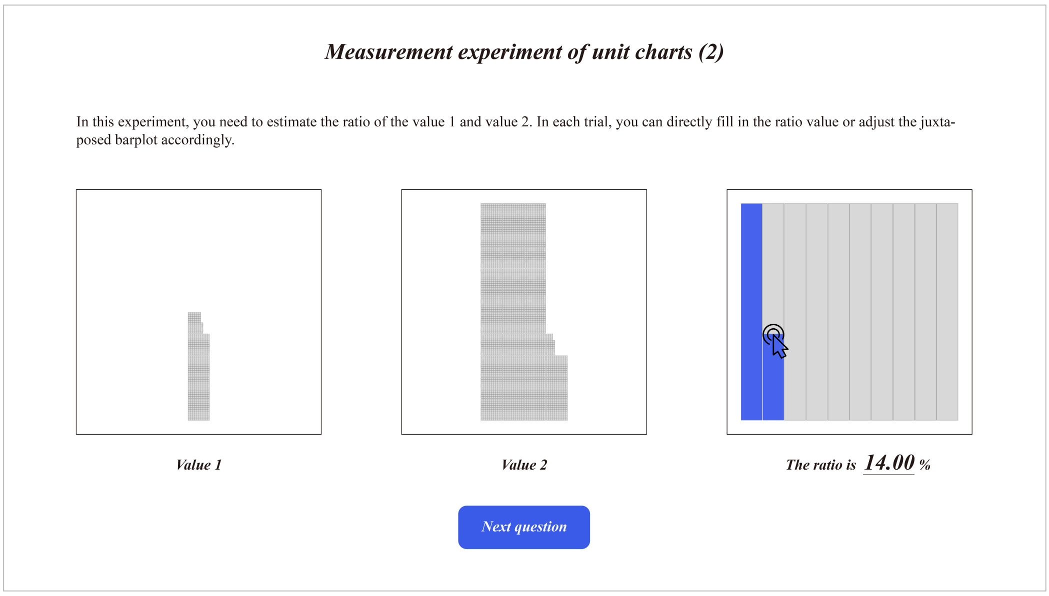


Figure 9. A screenshot of the measurement experiment for unit charts.

4.5. Results

A total of 810 sets of user performance data were collected for Task 1, including correct answers, response answers, and response time. We calculated the error rate of participants' responses in each data set. Figure 10a,b showed box plots for error rate and response time; outliers were omitted. One-way ANOVA was used to analyze the effect of different orders of magnitude on the two dependent variables. If approximate, the continuous variables between the two groups were compared with the student's *t*-test or Tamhane test. As shown in Figure 10a, most of the error rate was concentrated at 0, resulting in most of the data not being visible in the box plot. We tried logarithmic or square root transformation, but this still did not improve the visibility of the error rate data. The line chart in Figure 10a indicated the number of incorrect answers submitted by participants. It can be seen that as the order of magnitude increases, the number of errors made by participants becomes greater. However, this study found no significant difference in error rate when participants counted different order-of-magnitude values, F (3, 356) = 2.147, $p = 0.094$. There was a significant difference in the effect of the order of magnitude on response time. Tamhane test showed that participants had the longest response time when unit charts representing thousands digits (M = 34.370 s), followed by hundreds digits (M = 24.045 s), then tens digits (M = 14.394 s), and the shortest was single digits (M = 8.098 s). Table 2 summarizes the results of multiple comparisons.

Task 2 collected 1620 sets of user performance data, each set including raw stimulus data, stimulus proportion, adjustment proportion, and response time. We calculated the error rate of participants' adjustment in each data set. A two-way ANOVA was used to analyze the effects of the base and the Diss_OM on error rate and response time. This study used box plots to identify outliers, Shapiro–Wilk tests to determine data normality, and Levene chi-square tests to determine the equal variance. Figure 11a,b showed the results of descriptive statistics for error rate and response time, respectively.

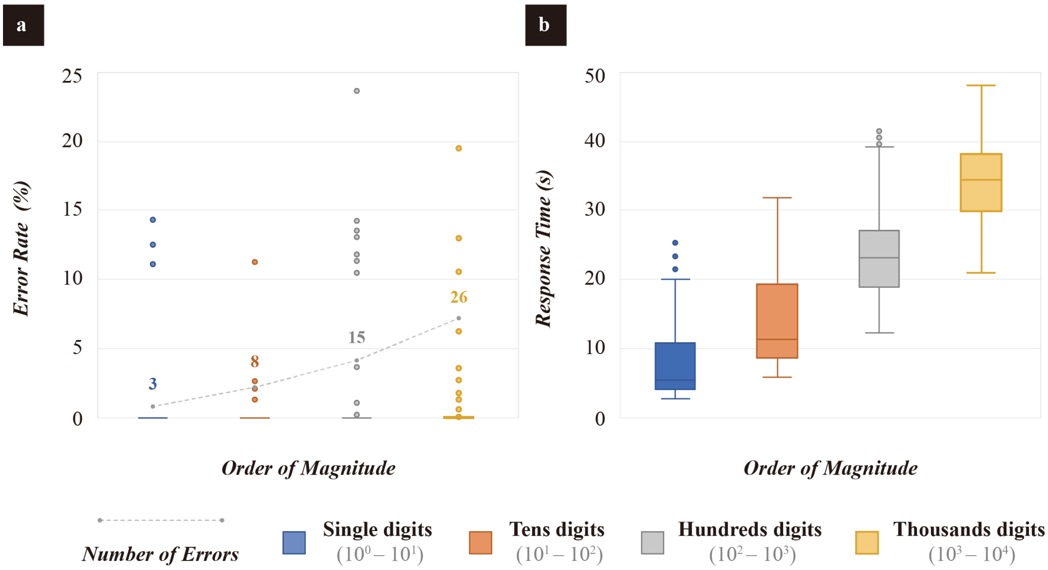


Figure 10. Box plot of error rate and response time for unit chart representing different orders of magnitude. (**a**) Descriptive statistics for Error rate. (**b**) Descriptive statistics for response time.

Table 2. Post hoc multiple comparisons with the Tamhane test for response time.

(I) OM	(J) OM	Mean Difference (I – J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Single	Tens	−6.297 *	0.920	0.000	−8.746	−3.847
	Hundreds	−15.948 *	0.915	0.000	−18.383	−13.513
	Thousands	−26.272 *	0.836	0.000	−28.496	−24.048
Tens	Hundreds	−9.651 *	1.010	0.000	−12.337	−6.965
	Thousands	−19.975 *	0.939	0.000	−22.474	−17.477
Hundreds	Thousands	−10.324 *	0.933	0.000	−12.808	−7.840

*. The mean difference is significant at the 0.05 level.

There was no interaction between the base and the Diss_OM on error rate, F (2, 534) = 0.008, $p = 0.992$. There was no significant difference in the main effects among the base 8 and 10 ($p = 0.930$). The main effects analysis suggested a statistically significant effect of Diss_OM on error rate F (2, 534) = 61.233, $p < 0.001$. The results of multiple comparisons were analyzed by a Tukey HSD test, as shown in Table 3. Pairwise comparisons were used to analyze the main effects results of the Diss_OM. Among them, the error rate of participants estimating the unit charts with the Diss_OM of 3 was 28.23% higher than that of 1 (95% CI 22.04–34.43%), $p < 0.001$. The estimation error rate for unit charts with the Diss_OM of 3 was 20.49% higher than that of 2 (95% CI 14.29–26.69%), $p < 0.001$. The error rate of estimation when the Diss_OM were 2 was 7.74% higher than when those were 1, (95% CI 1.55–13.94%), $p < 0.05$.

In the analysis of response time, the base and the Diss_OM interacted with F (2, 534) = 7.003, $p < 0.05$. Simple main effects analysis suggested that participants' response time was significantly different when estimating charts with different Diss_OM F (2, 534) = 770.717, $p < 0.01$, whereas no significant differences were detected for unit charts with different bases F (1, 534) = 3.799, $p = 0.052$. Analysis of variance with a Tukey HSD test was performed for multiple comparisons, as shown in Table 4. The multiple comparisons showed that participants took 5.950 s longer to estimate unit charts with Diss_OM of 3 than those with that of 1 (95% CI 4.764–7.136), $p < 0.001$. The estimation time for unit charts with Diss_OM of 3 was 2.282 s longer than that of 2 (95% CI 1.096–3.469), $p < 0.001$. There was a 2.282-second difference between unit charts with Diss_OM of 3 and 2 (95% CI 1.096–3.469), $p < 0.001$.

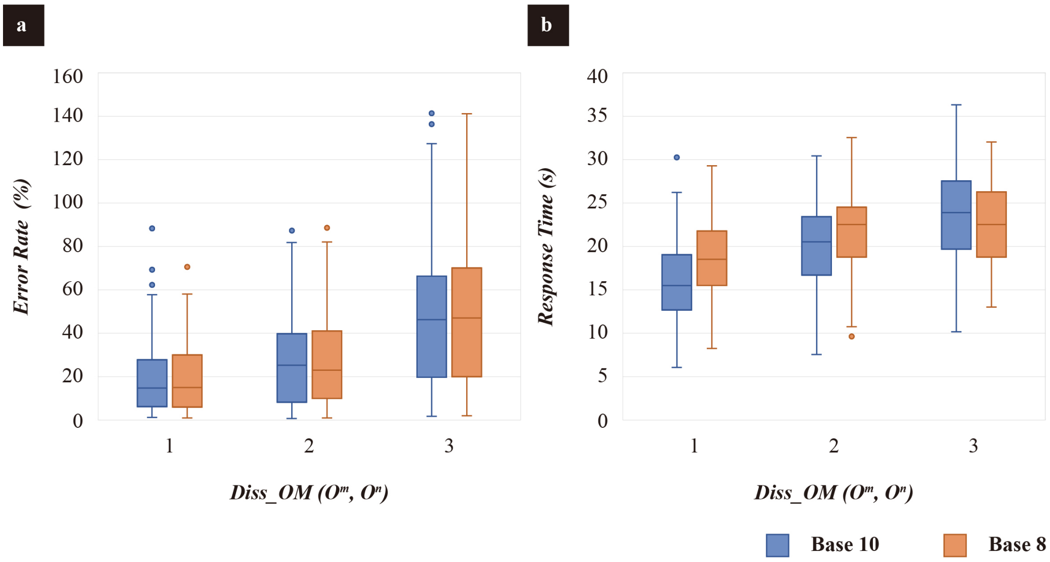


Figure 11. Box plot of error rate and response time for unit chart representing different orders of magnitude (**a**) Descriptive statistics for Error rate. (**b**) Descriptive statistics for response time.

Table 3. Post hoc multiple comparisons with the Tukey HSD test for *JND*.

(I) OM	(J) OM	Mean Difference (I − J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	−0.077 *	0.016	0.009	−0.139	−0.344
	3	−2.823 *	0.026	0.000	−0.016	−0.221
2	3	−0.205 *	0.027	0.000	−0.267	−0.143

*. The mean difference is significant at the 0.05 level.

Table 4. Post hoc multiple comparisons with the Tukey HSD test for *JND*.

(I) OM	(J) OM	Mean Difference (I − J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	−3.668 *	0.505	0.000	−4.854	−2.481
	3	−5.950 *	0.516	0.000	−7.136	−4.763
2	3	−2.282 *	0.512	0.000	−3.4695	−1.096

*. The mean difference is significant at the 0.05 level.

5. Discussions

This study conducted a series of measurement studies for unit charts in the context of large-order-of-magnitude data visualization. As opposed to the previous studies [2,5,6], this study examined the potential costs or benefits of unit charts when used to characterize perceptual features of large-scale data rather than visual expressiveness. The two experiments in this study yielded several results that can be used to guide future research.

First, in terms of graphical perception mechanisms, participants could slowly count the exact value of unit charts for values up to 10,000. Participants counted each magnitude in turn, and the overall error would remain low if the largest magnitude were not miscounted. However, when it came to comparing the proportions of the two charts, participants switched to noisy estimation mode. We also found that participants would use visual anchors for noisy estimation, which was consistent with the finding of Schiano and Tversky [56]. Cleveland and McGill claimed that bar charts are fairly accurate as they use length or position estimation (depending on configuration), and area estimation is typically less accurate [55]. However, it is worth noting that unit charts showed more advantages because they have better discriminability specifically in the case of comparisons of values across multiple orders of magnitude, as mentioned in the introduction.

Second, both Experiments 1 and 2 showed that the response time increases as the magnitude of the characterization increases. The limited working memory may explain this — the larger the magnitude, the more entries are cached in the processing memory, and thus larger values require more memory resources [2]. This view is corroborated by previous research findings [32,33], such as Landy, Silbert, and Goldin noted that even adults may have difficulties in estimating the numbers on larger ranges (1000–1 billion) number lines [33].

Finally, in terms of perceptual discriminability, this experiment's theoretical limit of human eye strength can be derived from the angle of view (AOV) calculation Equation (1) as 0.76 pixels. However, our results showed that the perceptual discriminability of a unit chart is not fixed despite the fixed distance between the participant's eyes and the screen. As the complexity of the graph increases, a greater visual distance is required for the viewer to identify the difference accurately. Hence, we recommend appropriately scaling up unit graphs size when representing multiscale data sets. Interestingly, while Experiment 1 showed that *JND* increases with numeric values, they fail to fit significantly according to the Weber-Fechner law [58]. One possible explanation might be that the maximum edge length in this experiment is limited to 11 px, and the adaptive algorithm encounters the floor effect of the *JND* from above. Another possible explanation is the unit charts' particular representation pattern—representing a value by multiples. When unit charts reach an integer order of magnitude, the complete morphology acts as an additional cue, thus acting as a redundancy and reducing messaging errors and noise. We suspect that there is additionally a separation effect [27]—the amount of distance between units is also important in determining intensity. Obviously, the Weber-Fechner relationship fails at low and very high intensities, but these are still something worth addressing.

$$\frac{\theta}{2} = \arctan \frac{H}{2D}^{*} \tag{1}$$

* For this experiment, a 23.8-inch monitor (539.1 × 321.1 mm) with 1920 × 1080 pixels resolution was used, and the viewers' eye distance to the screen "D" was 70 cm.

6. Conclusions and Future Work

This study aimed to guide the designing of unit charts by investigating the perceptual characteristics of unit charts in the context of representing asymmetric interactions such as large-scale numbers. Our results indicated that the order of magnitude of the values impacts both the *JND* and response time when counting unit charts. Specifically, when performing slow counting, the *JND* and the response time viewers required increased due to the complexity of the unit charts representing larger values. However, no significant differences in accuracy were detected. A difference was observed between noisy estimation and slow counting results when performing a magnitude comparison task. When the Diss_OM between the compared values were larger, the response time grew, and notably, the estimation error increased as well. It is worth noting that these two findings may not apply to magnitude estimates in more extensive ranges above 100,000. This study supports and complements previous research and may contribute to understanding how viewers perceive unit charts and the factors that influence perception. Compared with traditional charts (e.g., bar charts, pie charts, etc.), unit charts perform well when providing an overview of data containing multiple orders of magnitude. Viewers can still obtain relatively accurate scaling relationships between individual items using unit charts, even with limited screen size and resolution. When representing more extensive ranges of values above 100,000, more research is needed to determine how people reason about magnitudes beyond human perception. Future work will focus on the perception of values beyond 10,000 and different levels of user training and familiarity.

Author Contributions: Conceptualization, Y.L.; methodology, Y.L. and W.T.; software, W.T.; validation, Y.L. and F.S.; formal analysis, Y.L. and Y.Z.; investigation, F.S.; resources, Y.T.; data curation, Y.Z.; writing—original draft preparation, Y.L.; writing—review and editing, Y.L. and W.T.; visualization, Y.L. and W.T.; supervision, Y.T.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: The paper is supported jointly by the New Liberal Arts Research and Reform Practice Project (No. 2021160033), the Special Project of Nanjing University of Science and Technology for Independent Research-Cross-disciplinary Cultivation (No. XJ2022000102), and the new liberal arts education reform project of the Ministry of Education (No. 21YJC760072).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data available on request due to restrictions (e.g., privacy or ethical reasons). The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy reasons.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

JND	just-noticeable difference
AOV	the angle of view
OM	the Order of Magnitude
Diss_OM	the difference in the order of magnitude

References

1. Amini, F.; Riche, N.H.; Lee, B.; Leboe-McGowan, J.; Irani, P. Hooked on Data Videos: Assessing the Effect of Animation and Pictographs on Viewer Engagement. In Proceedings of the 2018 International Conference on Advanced Visual Interfaces, Castiglione della Pescaia Grosseto, Italy, 29 May–1 June 2018.
2. Haroz, S.; Kosara, R.; Franconeri, S. Isotype Visualization: Memory, Performance, and Engagement. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea, 18–23 April 2015.
3. Romat, H.; Henry Riche, N.; Hurter, C.; Drucker, S.; Amini, F.; Hinckley, K. Dear Pictograph: Investigating the Role of Personalization and Immersion for Consuming and Enjoying Visualizations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020.
4. Setlur, V.; Mackinlay, J.D. Automatic Generation of Semantic Icon Encodings for Visualizations. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, ON, Canada, 26 April—1 May 2014.
5. Garcia-Retamero, R.; Cokely, E.T. Communicating Health Risks with Visual Aids. *Curr. Dir. Psychol. Sci.* **2013**, *22*, 392–399. [CrossRef]
6. Borkin, M.A.; Vo, A.A.; Bylinskii, Z.; Isola, P.; Sunkavalli, S.; Oliva, A.; Pfister, H. What Makes a Visualization Memorable? *IEEE Trans. Vis. Comput. Graph.* **2013**, *19*, 2306–2315. [CrossRef] [PubMed]
7. Harris, R.L. *Information Graphics: A Comprehensive Illustrated Reference*; Oxford University Press: Oxford, UK, 2000; pp. 154–196.
8. Neurath, O.T.T.O. *From Hieroglyphics to Isotype: A Visual Autobiography*; PERPENSA PRESS: London, United Kingdom, 2019; pp. 154–196.
9. Bresnahan, K. "An Unused Esperanto": Internationalism and Pictographic Design, 1930–70. *Des. Cult.* **2011**, *3*, 5–24. [CrossRef]
10. Fagerlin, A.; Wang, C.; Ubel, P.A. Reducing the Influence of Anecdotal Reasoning on People's Health Care Decisions: Is a Picture Worth a Thousand Statistics? *Med. Decis. Mak.* **2005**, *25*, 398–405. [CrossRef] [PubMed]
11. McCaffery, K.J.; Dixon, A.; Hayen, A.; Jansen, J.; Smith, S.; Simpson, J.M. The Influence of Graphic Display Format on the Interpretations of Quantitative Risk Information among Adults with Lower Education and Literacy. *Med. Decis. Mak.* **2011**, *32*, 532–544. [CrossRef]
12. Boy, J.; Pandey, A.V.; Emerson, J.; Satterthwaite, M.; Nov, O.; Bertini, E. Showing People behind Data. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017.
13. Morais, L.; Jansen, Y.; Andrade, N.; Dragicevic, P. Showing Data about People: A Design Space of Anthropographics. *IEEE Trans. Vis. Comput. Graph.* **2022**, *28*, 1661–1679. [CrossRef]
14. Burns, A.; Xiong, C.; Franconeri, S.; Cairo, A.; Mahyar, N. Designing with Pictographs: Envision Topics without Sacrificing Understanding. *IEEE Trans. Vis. Comput. Graph.* **2021**, *28*, 4515–4530. [CrossRef]

15. Wang, N.; Burns, L. Visual Design Thinking for Public Education: Improving Knowledge Development and Exchange via Transformative PICTOGRAM DESIGN. In Proceedings of the 2020 24th International Conference Information Visualisation (IV), Melbourne, Australia, 7–11 September 2020.
16. Schreder, G.; Hynek, N.; Windhager, F.; Mayr, E. Rediscovering Isotype from a Cognitive Perspective. In *Diagrammatic Represent. Inference*; Springer: Cham, Switzerland, 2018; pp. 458–469.
17. Tufte, E.R. *The Visual Display of Quantitative Information*; Graphics Press: Cheshire, CT, USA, 2018; pp. 154–196.
18. Unit Charts Are For Kids. Available online: http://www.perceptualedge.com/articles/visual_business_intelligence/unit_charts_are_for_kids.pdf (accessed on 11 December 2022).
19. Pino, T. Counting But Losing Count: The Legacy of Otto Neurath's Isotype Charts. *Visible Lang.* **2018**, *52*, 83–109.
20. Cakmak, E.; Jackle, D.; Schreck, T.; Keim, D.; Fuchs, J. Multiscale Visualization: A Structured Literature Analysis. *IEEE Trans. Vis. Comput. Graph.* **2021**, *28*, 4918–4929. [CrossRef]
21. Randal, M. Radiation. Available online: https://www.explainxkcd.com/wiki/index.php/Radiation (accessed on 11 December 2022).
22. Keim, D.A. Information Visualization and Visual Data Mining. *IEEE Trans. Vis. Comput. Graph.* **2002**, *8*, 1–8. [CrossRef]
23. Tang, J.; Liu, J.; Zhang, M.; Mei, Q. Visualizing Large-Scale and High-Dimensional Data. In Proceedings of the 25th International Conference on World Wide Web, Montreal, QC, Canada, 11–15 April 2016.
24. Hlawatsch, M.; Sadlo, F.; Burch, M.; Weiskopf, D. Scale-Stack Bar Charts. *Comput. Graph. Forum* **2013**, *32*, 181–190. [CrossRef]
25. Wünsche, B. A Survey, Classification and Analysis of Perceptual Concepts and Their Application for the Effective Visualisation of Complex Information. In Proceedings of the APVis '04, 2004 Australasian symposium on Information Visualisation, Christchurch, New Zealand, 1 February 2004; Volume 35, pp. 17–24.
26. Moody, D. The "Physics" of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering. *IEEE Trans. Softw. Eng.* **2009**, *35*, 756–779. [CrossRef]
27. Garner, W.R. *The Processing of Information and Structure*; Psychology Press: New York, NY, USA, 2009; pp. 154–196.
28. Resnick, I.; Newcombe, N.S.; Shipley, T.F. Dealing with Big Numbers: Representation and Understanding of Magnitudes Outside of Human Experience. *Cogn. Sci.* **2016**, *41*, 1020–1041. [CrossRef] [PubMed]
29. Choo, H.; Franconeri, S.L. Enumeration of Small Collections Violates Weber's Law. *Psychon. Bull. Rev.* **2013**, *21*, 93–99. [CrossRef]
30. Haroz, S.; Whitney, D. How Capacity Limits of Attention Influence Information Visualization Effectiveness. *IEEE Trans. Vis. Comput. Graph.* **2012**, *18*, 2402–2410. [CrossRef] [PubMed]
31. Kaufman, E.L.; Lord, M.W.; Reese, T.W.; Volkmann, J. The Discrimination of Visual Number. *Am. J. Psychol.* **1949**, *62*, 498–525. [CrossRef] [PubMed]
32. Thompson, C.A.; Opfer, J.E. How 15 Hundred Is like 15 Cherries: Effect of Progressive Alignment on Representational Changes in Numerical Cognition. *Child Dev.* **2010**, *81*, 1768–1786. [CrossRef]
33. Landy, D.; Silbert, N.; Goldin, A. Estimating Large Numbers. *Cogn. Sci.* **2013**, *37*, 775–799. [CrossRef] [PubMed]
34. Resnick, I.; Davatzes, A.; Newcombe, N.S.; Shipley, T.F. Using Relational Reasoning to Learn about Scientific Phenomena at Unfamiliar Scales. *Educ. Psychol. Rev.* **2016**, *29*, 11–25. [CrossRef]
35. Bueti, D.; Walsh, V. The Parietal Cortex and the Representation of Time, Space, Number and Other Magnitudes. *Philos. Trans. R. Soc. Biol. Sci.* **2009**, *364*, 1831–1840. [CrossRef]
36. Cantlon, J.F.; Platt, M.L.; Brannon, E.M. Beyond the Number Domain. *Trends Cogn. Sci.* **2009**, *13*, 83–91. [CrossRef] [PubMed]
37. Lourenco, S.F.; Longo, M.R. Origins and Development of Generalized Magnitude Representation. In *Space, Time and Number in the Brain*; Elsevier: Amsterdam, The Netherlands, 2011; pp. 225–244.
38. Walsh, V. A Theory of Magnitude: Common Cortical Metrics of Time, Space and Quantity. *Trends Cogn. Sci.* **2003**, *7*, 483–488. [CrossRef] [PubMed]
39. Cohenkadosh, R.; Lammertyn, J.; Izard, V. Are Numbers Special? an Overview of Chronometric, Neuroimaging, Developmental and Comparative Studies of Magnitude Representation. *Prog. Neurobiol.* **2008**, *84*, 132–147. [CrossRef] [PubMed]
40. Dehaene, S.; Bossini, S.; Giraux, P. The Mental Representation of Parity and Number Magnitude. *J. Exp. Psychol. Gen.* **1993**, *122*, 371–396. [CrossRef]
41. Ishihara, M.; Keller, P.; Rossetti, Y.; Prinz, W. Horizontal Spatial Representations of Time: Evidence for the STEARC Effect. *Cortex* **2008**, *44*, 454–461. [CrossRef]
42. Moyer, R.S.; Landauer, T.K. Time Required for Judgements of Numerical Inequality. *Nature* **1967**, *215*, 1519–1520. [CrossRef]
43. Hubbard, E.M.; Piazza, M.; Pinel, P.; Dehaene, S. Interactions between Number and Space in Parietal Cortex. *Nat. Rev. Neurosci.* **2005**, *6*, 435–448. [CrossRef]
44. Izard, V.; Dehaene, S. Calibrating the Mental Number Line. *Cognition* **2008**, *106*, 1221–1247. [CrossRef]
45. Barth, H.C.; Paladino, A.M. The Development of Numerical Estimation: Evidence against a Representational Shift. *Dev. Sci.* **2010**, *14*, 125–135. [CrossRef]
46. Opfer, J.E.; Siegler, R.S.; Young, C.J. The Powers of Noise-Fitting: Reply to Barth and Paladino. *Dev. Sci.* **2011**, *14*, 1194–1204. [CrossRef]
47. Dehaene, S.; Izard, V.; Spelke, E.; Pica, P. Log or Linear? Distinct Intuitions of the Number Scale in Western and Amazonian Indigene Cultures. *Science* **2008**, *320*, 1217–1220. [CrossRef] [PubMed]

48. Dehaene, S. The Neural Basis of the Weber–Fechner Law: A Logarithmic Mental Number Line. *Trends Cogn. Sci.* **2003**, *7*, 145–147. [CrossRef] [PubMed]
49. Shipley, T.F.; Zacks, J.M. *Understanding Events: From Perception to Action*; Oxford University Press: Oxford, UK, 2008; pp. 154–196.
50. Zacks, J.M.; Tversky, B. Event Structure in Perception and Conception. *Psychol. Bull.* **2001**, *127*, 3–21. [CrossRef] [PubMed]
51. Nutter, J.; Forrest, W. Weber–Fechner Law. *Encycl. Res. Des.* **2010**, *3*, 1612–1615.
52. Dobbert, T. *Matchmoving: The Invisible Art of Camera Tracking*; PWiley: Indianapolis, IN, USA, 2013; pp. 154–196.
53. Elliott, M.A.; Nothelfer, C.; Xiong, C.; Szafir, D.A. A Design Space of Vision Science Methods for Visualization Research. *IEEE Trans. Vis. Comput. Graph.* **2021**, *27*, 1117–1127. [CrossRef]
54. SPSS in Windows: ANOVA. Available online: https://www.mdthinducollege.org/ebooks/statistics/anova/SPSS_20ANOVA.pdf (accessed on 11 December 2022).
55. Cleveland, W.S.; McGill, R. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *J. Am. Stat. Assoc.* **1984**, *79*, 531–554. [CrossRef]
56. Schiano, D.J.; Tversky, B. Structure and Strategy in Encoding Simplified Graphs. *Mem. Cogn.* **1992**, *20*, 12–20. [CrossRef]
57. Saket, B.; Srinivasan, A.; Ragan, E.D.; Endert, A. Evaluating Interactive Graphical Encodings for Data Visualization. *IEEE Trans. Vis. Comput. Graph.* **2018**, *24*, 1316–1330. [CrossRef]
58. Kay, M.; Heer, J. Beyond Weber's Law: A Second Look at Ranking Visualizations of Correlation. *IEEE Trans. Vis. Comput. Graph.* **2016**, *22*, 469–478. [CrossRef]