Correlational Data Visualizations with Colored Bar Charts

Bdour Alwuqaysi
King's College London
bdour.alwuqaysi@kcl.ac.uk

Alfie Abdul-Rahman
King's College London
alfie.abdulrahman@kcl.ac.uk

Rita Borgo
King's College London
rita.borgo@kcl.ac.uk

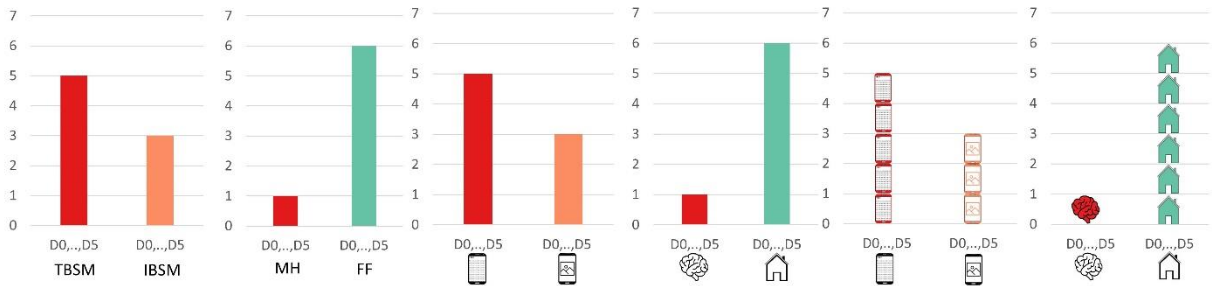

Figure 1. Examples of study stimuli that show variations of colored vertical bar charts.

ABSTRACT

This study evaluates the efficacy of three varied designs of colored bar charts—featuring text annotations, icon shapes, and stacked icons—in illustrating the correlation between social media use, mental health, and family functioning. It seeks to bridge a research gap by exploring this specific aspect of visualization techniques. The study uses an online survey with 500 participants to evaluate these designs in terms of accuracy, response time, and user preferences. The research findings suggest that colored bar charts with text can reduce response time in medium complexity tasks. Meanwhile, colored bar charts with text and stacked icons can enhance accuracy in medium and hard complexity tasks. Colored bar charts with icons can engage users more in medium and hard tasks, and consistently demonstrate high interaction. The study revealed that gender could influence response time and interaction, with colored bar charts with icons generally preferred. However, visualization preferences may vary across age groups, highlighting the importance of personalized visualizations for diverse users. These insights are crucial for individuals aiming to utilize visualizations effectively for correlational data, prompting further investigation into enhancing data visualization through subsequent research based on these results.

Keywords

Correlational data visualizations, Colored bar charts, Text, Icons, Stacked icons.

1. INTRODUCTION AND BACKGROUND

In the domain of data visualization, the design of visual representations significantly influences how users interpret and engage with data. Previous research has underscored the importance of intuitive and visually appealing visualizations in facilitating comprehension and decision-making [1, 2, 3]. However, the relationship between visualization complexity and user behavior remains complex and multifaceted. Correlational data representation studies offer a unique opportunity to explore this relationship by examining how variations in visualization design correlate with user responses. Understanding these correlations is essential for creating visualizations that effectively convey information while minimizing cognitive load and enhancing user engagement. Moreover, demographic factors such as gender, and age can further influence user preferences and interactions with visualizations. Studies have shown that individuals from different demographic backgrounds may have varying cognitive styles and preferences when interacting with visual information [4, 5]. Therefore, exploring how demographic factors intersect with visualization design is crucial for ensuring the inclusivity and effectiveness of visualizations across diverse user groups. By focusing on correlational data, this study seeks to uncover relationships between visualization complexity, user behavior, and demographic characteristics. In recent years, visualization techniques has become increasingly important in improving user comprehension and engagement with data. Among these techniques, colored bar charts with text, colored bar charts with icons, and colored stacked icons with icons have emerged as methods for conveying information. The effectiveness of some of these visualization techniques in facilitating user understanding and engagement has been demonstrated in various studies [6, 7, 8]. However, the impact of these techniques on user preferences and interactions, particularly in correlation with demographic factors such as gender, and age remains an area of exploration. Thus, this study aims to explore the relationship between correlational data visualization designs and user responses regarding demographic characteristics.

Contributions:

The study revealed several key findings. Firstly, a colored bar chart with text was more effective in reducing response time than a colored bar chart with icons, especially in medium complexity designs. Secondly, a colored bar chart with text and colored stacked icons showed greater accuracy compared to a colored bar chart with icons in both medium and hard complexities. Thirdly, the colored bar chart with icons was the most engaging and efficient design for completing tasks. Particularly in medium and hard complexities, with increasing repetitions, response time decreased across all design complexities. The colored bar chart with icons consistently shows higher interaction levels, highlighting the potential to enhance user engagement and learning efficiency. Furthermore, the study found gender variations in the responses. Males exhibited lower response times in certain designs. Preferences also varied among participants. The colored bar chart with icons was the most preferred due to its appeal and understandability. In contrast, the colored stacked icons with icons were favored for their memorability and potential for triggering behavioral changes. The majority favored the colored bar chart with icons. Lastly, younger participants tended to prefer the colored bar chart with icons, while older individuals favored the colored bar chart with text.

2. METHODOLOGY

An online study was conducted to investigate the topic. Participants have started with detailed training. This was followed by the main segment of the study,

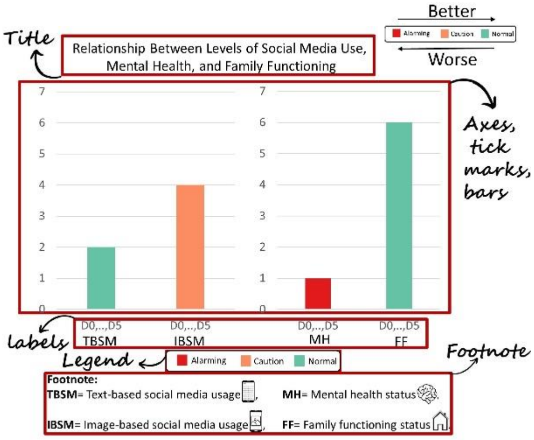

Figure 2. Training was provided to participants to help them understand each visual encoding.

consisting of stimuli and questions prompting individuals to conclude from the presented visualizations. Inspired by the current methodology in vision science [9], we deliberately selected methods employing a familiar chart type (bar chart) as the foundation for stimuli design and tasks to enhance relevance, simulating common data visualization interactions. The process began with creating various visualizations to represent our topic-specific correlational data design, emphasizing having these visualizations evaluated by the participants. Efficient data visualization (i.e., high accuracy and short response time) requires careful comparison of encodings and layouts, accurate value estimation, magnitude sense, and subdivision within charts [10]. Our core stimuli tasks were visual search, and comparison. We also combined data from two charts with inverted color scales.[1]

2.1 Study Structure

The online study was conducted using the Qualtrics platform [11], Participants received a comprehensive information sheet and consent form explicitly stating their right to withdraw from the study at any point. A computer screen was a prerequisite for participation, and individuals were required to have normal or corrected-to-normal vision without any color-vision deficiencies. At the outset, two questions utilizing the Ishihara test [12], a widely recognized test in visualization studies involving colors, were employed to identify red-green color deficiencies among participants. To ensure the accuracy of the results, we excluded responses to the two color-deficiency screening questions from subsequent data analysis. An additional attention check question was used to verify the survey results, but it was not included in the data analysis. No personally identifiable information was collected, and all data were securely stored on a

[1] Refer to the supplementary material for additional details: https://anonymous.4open.science/r/Vis-5C01/Qualtrics%20survey%20A.pdf

private server at the author's institution. The study comprised three main sections. Participants were asked seven demographic questions followed by three questions about their social media usage after signing the consent form. In the study's second section, introduced. To ensure participants understood the concepts, definitions of key terms were provided customized training and visualizations were along with a sample question they needed to answer correctly before moving on. The training helped participants understand each visual encoding by presenting charts, annotations, and brief textual explanations, i.e., The participants were provided with to clarify the meaning of 'Better' and 'Worse' in addition to the definition of these concepts in the context of the study (see Fig.2). Although the questions from the training phase were not included in the data analysis, they were used as a quality assurance measure of the results. A clear separation was added between the training and visualization sections to prepare and prompt the participants. The visualization section included three design variations presented at three complexity levels (easy, medium, and hard), with two repetitions to validate results. Participants confirmed their readiness to proceed to the next block, allowing breaks as needed. Response time (TSubmit), interaction-user engagement (Click Count), and accuracy (Percentage of Wrong Answers) were assessed for each visualization. The participants were asked to choose the visualizations based on their understandability, appeal, memorability, and identification of the most useful visual design feature when answering questions. Two levels of randomization were used to increase result validity: one within each section's complexity and another within each question's answer choices.

Screening: Participants were started with seven demographic questions to gather information about gender, age, ethnicity, education level, employment status, field of expertise, and any prior experience in visualization. Only participants with normal or corrected-to-normal vision without any color-vision deficiencies were included to ensure a consistent and valid sample. We do this by asking participant to correctly answer the two Ishihara test questions to be included in the analysis. Participation in the study was restricted to individuals using a laptop or desktop computer with an internet connection, facilitating access to the web browser required for the study. Data from 39 participants were excluded due to failure to pass attention checks, missing responses, or experiencing a server interruption during the study. The data collection process continued until a high-quality dataset was achieved (i.e., characterized by a balanced sample across genders).

2.1.1. Stimuli Design: Visual stimuli for this online study were created using draw.io, a cross-platform

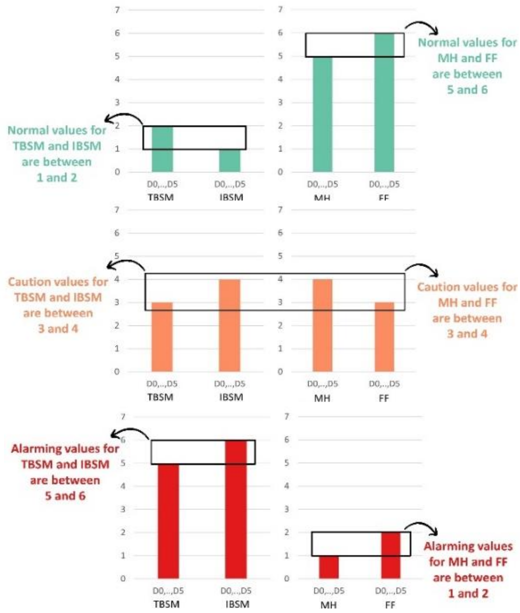

Figure 3. The color scale was used in this study.

graph drawing software developed in HTML5 and JavaScript. Stimuli were saved as static images and presented to participants using the Qualtrics platform. Visual task performance was tested over the four manipulation elements in the stimuli: data density, chart type, color scale, and task difficulty.

Data: The data used in the visualizations of this study was real-world data collected from an Android application that the researcher had previously developed in 2023. This application was used to collect self-reported data on mental health and family functioning, and track social media usage for a month. The data has already been analyzed and proven correlational [13], making it the wel-suited dataset for this study. This study uses a bar chart approach to display two sets of data side-by-side with different color scales. The left bar chart (LC) depicts the correlational data of individuals' text-based social media use (TBSM) and image-based social media use (IBSM) for six days, while the right bar chart (RC) displays the anonymous records of mental health status (MH) and family functioning status (FF) for the same six days (see Fig.1). The data in our study are represented by integer numbers floating points are rounded to the nearest integer value. The data is presented at three different levels of density, with two, four, and six points on each chart, making a total of 4, 8, and 12 points across the two bar charts in each correlational data representation (see Fig.4). Each presented stimulus was unique, representing a new data record without repetition or manipulation. The human mind only holds around 7 points in immediate memory and absolute judgment. Going over this limit increases cognitive load [14]. Therefore, we chose set sizes of 8 and 12 for our subsequent data points. To align with the left-to-right reading sequence of English readers, we placed social media use charts on the left

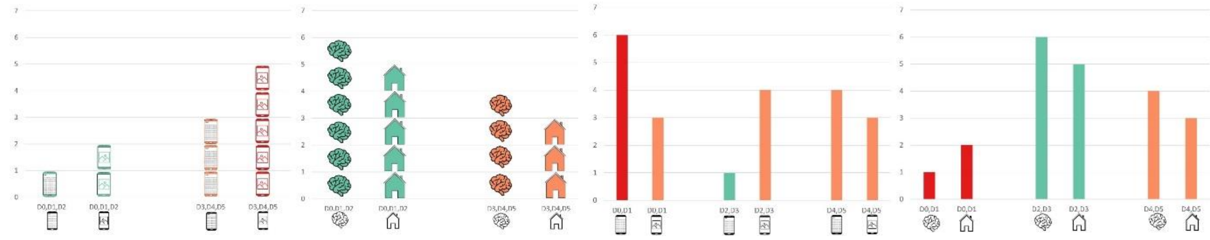

Figure 4. Stimulus density level 8 & 12 examples.

to create a logical flow. This arrangement establishes a clear narrative progression by introducing social media variables before revealing their impact on mental health and family functioning. This approach helps readers understand the correlational data represented. The LC begins with TBSM and progresses to IBSM, following a logical and chronological sequence. This mirrors the provided information and emphasizes the foundational nature of TBSM. On the RC, MH precedes FF, emphasizing that personal mental health serves as a precursor to the health of the family unit. Arranging similar variables together facilitates easy comparison, helping readers evaluate the impact of social media on mental health and family function. Aligning charts with corresponding social media types underscores potential causal relationships to enhance comprehension. We varied set sizes across charts while maintaining consistent density for side-by-side comparison. As detailed below, we also gathered visualization assessments, including selected answers, response times, and participant preferences of each chart type.

Chart Components: The charts in this study have the same title, axes, tick marks, legend, and footnote. Although they have similarities, the three designs differ in their use of bars and labels. Design I features colored bars with text, Design II uses colored bars with icons, and Design III displays colored stacked icons with icons, as shown in Fig.1 and Fig.2.

Chart Types: Design I colored bar charts with text offer a simple way to present data, combining visuals with annotations for context [15]. Design II colored bar charts with icons allow for quick identification of categories [16, 17]. Design III colored stacked icons with icons are used for complex datasets, allowing for easy interpretation of multidimensional data [18].

Color Scale: The color scales employed in our correlational data bars representation differ between the LC, representing TBSM and IBSM, and the RC, representing MH and FF for individuals (see Fig.3). The green bars on LC denote normal values between 1 and 2, where lower social media consumption is considered better, as proven in the literature [19]. In contrast, normal values on RC fall between 5 and 6, where higher scores of MH and FF represent the good state of individuals based on the literature [20, 21]. The orange color on both LC and RC indicates caution values between 3 and 4, displaying intermediate TBSM, IBSM, MH, and FF values. The red bars on LC signify alarming values between 5 and 6, where

increased social media consumption is considered worse, as proven by the literature [19]. Conversely, alarming values on RC are between 1 and 2, where lower scores of MH and FF represent the worst state of individuals, according to the literature [20, 21]. Inspired by the color theory principles [22], we designed the bars in different color variations to create a clear and effective visual display. Red is utilized to signify alarm and danger, while green represents normalcy and calmness. The orange is used to indicate caution and alertness. These color choices aim to elicit specific emotional responses and enhance the understanding of the information presented in the visualizations [23]. The shades of these colors have been selected from the Colors Brewer tool [24] to ensure friendliness to individuals with color blindness. Redundant color encoding used in this study can help differentiate data points more effectively. Combined with other visual variables, it creates unique visual patterns that aid pattern recognition and data segmentation. This improves visualization clarity, especially in dense contexts [22].

2.1.2. Task Design

The study used a primary task of visual search process and comparison, an essential element in interpreting visualizations. The search process task involved searching over two channels (i.e., two side-by-side charts with the same density but different color scales). The difficulty of this search process increased with the increase in object density. The study aimed to assess participants' performance in three different types of charts. There were three levels of density for each chart type.

Additionally, there were three multiple-choice questions with an alternative forced-choice format. The questions were designed to increase difficulty and compare the performance of the participants across varying levels of difficulty. These questions were categorized into Easy, Medium, and Hard, and the response options were randomized across all questions. In most questions, two answers are partially incorrect, while the third is entirely correct. This approach ensures the results are valid and accurate (refer to the supplementary material for more details).

Easy Question: The easy question involves a search and comparison task. It displays two side-by-side charts with values from the same set of data (i.e., (LC: TBSM, IBSM), (LR: MH, FF)) across 6 days. This question is classified as easy because it involves searching for specific target data. Once located, participants compare the target data across the charts. For instance, participants searched for IBSM in the chart and compared its value to the second set of data TBSM to determine if cases were Worse, Same, or Better over the same 6-day period. Example of an Easy Question: According to the above chart, is IBSM Worse, Same, or Better than TBSM? Easy Responses: IBSM is Worse than TBSM, IBSM is Same as TBSM, IBSM is Better than TBSM.

Medium Question: The medium question is a search and comparison task across an increased number of targets. It displays two side-by-side charts with values from two sets of data (i.e., (LC: TBSM, IBSM), (LR: MH, FF)) across 6 days. This question is classified as a medium as it increases the data set for search and comparison to two targets across both charts; participants search for the Worse, Same, or Better values and extract the category name and day for the Worse, Same, or Better value. Medium question example: Based on the above chart, which category and what days have the Better values? The options are provided to you in the format of (Category-days). Medium responses are (IBSM-D0,D1,D2) and (MH-D3,D4,D5), (TBSM-D3,D4,D5) and (FF-D0,D1,D2), (IBSM-D3,D4,D5) and (FF-D3,D4,D5).

Hard Question: The hard question is a search and comparison task that involves mathematical estimation. It displayed two side-by-side charts with values from two sets of data (i.e., (LC: TBSM, IBSM), (LR: MH, FF)) across 6 days. This question is classified as hard as after searching for the four target data sets across charts, participants had to perform compare them with mathematical symbols to estimate the answers. Participants must estimate the numerical Normal, Caution, and Alarming values in this question. Hard question example: In the chart above, what are the numerical Normal values for each individual column in different categories, where '&' represents the logical 'and' operation? Hard responses: TBSM & IBSM >6 and MH& FF >4, TBSM & IBSM <3 and MH& FF >4, TBSM & IBSM <3 and MH& FF >2.

2.1.3 Measures

We designed 27 multiple-choice questions by combining colored vertical bar charts on either right (RC) or left (LC), three levels of data density, and three question types/difficulty levels. To evaluate the performance, we recorded the chosen answer (Accuracy), response time (RT), and interaction (Click Count) for each question in the task parts of the study. We presented the stimuli randomly to participants to reduce any learning effects (e.g., bias in selection).

2.1.4 Preference

We used a two-phase question approach to determine the preferences of the participants. The first question asked the participants to evaluate various aspects of the designs presented, such as their favorite design, the one that was easiest to understand, the most visually appealing, the most memorable, and the potential for triggering behavioral change towards social media consumption. The participants had the option to choose from the provided designs or select "none." The second question aimed to identify the specific feature within the designs the participants relied on the most during the study. The options included colored bars with text, colored bars with icons, icons bar with icons, or an alternative option where participants could specify their choice. This approach provided a detailed understanding participant's self-reported preferences and the factors that influencing their decision-making process.

2.2 Recruitment

Participants were recruited through Prolific [25].To ensure the quality of our sample, participants underwent a pre-screening process, confirming normal vision and the absence of any deficiencies. Additionally, the Ishihara Color Test was employed as an extra measure to identify and exclude individuals with red-green color blindness, a common consideration in visualization studies [26]. The eligibility criteria included being 18 years and older, and fluency in English. We successfully achieved a balanced sample across genders, totaling 500 participants. Compensation adhered to Prolific's fair pay policy, with participants receiving £7.85 per hour. The average response time was 28 minutes and 39 seconds. A computer screen was explicitly required for participants to participate in the study. Data collection spanned six days, covering various time slots to enhance the diversity of potential participants. The study began on January 31, 2024, and concluded on February 5, 2024.

2.3 Participants

500 responses were collected, with approximately equal gender distribution (256 males, 252 females, and less than 1% others). The group of Gender “Others” has very low frequency (only 4); thus it is excluded from further statistical analysis. The age distribution reveals a significant representation in the 18-24 age group (23.05%) and a dominant presence in the 25-34 age group (49.02%). In this sample, individuals of Black ethnicity make up nearly a quarter (24.61%), while those of White ethnicity constitute about two-thirds (62.5%). Individuals with Mixed and Asian ethnicity make up around (5%). Arab ethnicity is less than (1%), and other ethnic groups comprise less than

(3%). Education-wise, the majority hold either a Bachelor's degree (42.77%) or a Master's (21.48%). Employment status indicates two-thirds are employed (50.2% full-time, 12.3% part-time, 8.79% self-employed), (15%) are students, and about (10%) are not employed. Participants span 27 fields, with the most prevalent being Business, Management, Marketing, and Related Support Services (15.23%), Computer and Information Sciences and Support Services (15.23%), Engineering Technologies and Engineering-Related Fields (10.74%), and Education (8.01%). Additionally, around (25%) of the population is distributed across various other fields.

2.4 Statistical Analysis

Python was used for most statistical analysis, except for Fisher's exact test and ranked two-way ANOVA, which were performed using R. Categorical variables were described by frequency and percentage, while numerical variables were presented as mean ± standard deviation, and median [minimum-maximum]. Before analysis, variables with low-frequency categories were either re-categorized or excluded. ANOVA was used to compare numerical variables among more than two groups. Welch ANOVA was used if homogeneity assumptions were violated, and Kruskal-Wallis test was applied if normality assumptions were violated. Pairwise comparison tests were performed for significant results. Two-way ANOVA on rank-transformed data was used to compare numerical variables among two grouping levels. A p-value less than 0.1 was considered significant in all tests.

2.5 Pilot Testing

We conducted a pilot test to identify potential refinements related to the length and complexity of the study. Based on the pilot results, specific questions in the study are reformatted for increased clarity and simplicity. As part of our efforts to enhance the overall understandability of the study's questions, we have included definitions for a few key concepts. These definitions are provided not only when the concepts are initially introduced but also when they are referenced in the questions.

2.6 Ethics

The user study described in this paper and other study that provided the data, have been approved by the Research Ethics Committee of the researcher institution (anonymized for submission).

3. RESULTS AND DISCUSSION

The evaluation results of all designs are summarized in Table 1 and elaborated further below. Our analysis has revealed that gender significantly affect response time and interaction. Moreover, age was significant factors affecting preferences, as evidenced by the p-values below.

3.1 Response Time

Designs Complexity: Participants' response times vary based on the complexity of the design. In easy designs, response times remain consistent across different designs. However, in medium designs, participants took less time on Design I (42.55 ± 72.92) than Design II (49.85 ± 59.86), showing a significant difference in TSubmit. Design III falls between Design I and II but is not significantly different from either. In hard designs, Design II and III (44.55 ± 40.71 and 44.0 ± 69.3 respectively) have significantly higher TSubmit than Design I (36.85 ± 42.16). These findings offer insights into how the complexity of task design influences response times. The findings regarding response time shed light on how design complexity impacts participants' performance. Consistent with previous literature, easy designs yielded steady response times across different designs, indicating minimal cognitive load [27].

Conversely, medium and hard designs demonstrated significant variations in response time, aligning with studies emphasizing the influence of task complexity on cognitive processing [28]. The observed differences between Design I and Design II in medium and hard designs suggest that certain design elements may pose greater cognitive challenges, resonating with prior research on the impact of design features on user response [29]. These insights underscore the importance of considering design complexity in visualization to optimize user efficiency and task performance.

Table 1. Evaluation factors across design complexity performance.

Evaluation Factors	Design Complexity Performance		
Measures	**Best**	**Intermediate**	**Worst**
Response time	I	III	II
Interaction	II	I	III
Accuracy	I	III	II
Self-reported			
General Preference	II	I	III
Understandability	II	I	III
Appeal	II	III	I
Memorability	III	II	I
May Trigger Behavioral Changes	III	II	I

Repetition: There is a significant decrease in TSubmit across all designs as repetitions increase, indicating participants' learning and familiarity with the target task. In easy designs, the median TSubmit in the last repetition (R3) (11.32 [1.83:753.34]) is nearly half of that in the first repetition (R1) (20.36 [3.45:1464.75]). In medium designs, R3 median TSubmit (20.25 [1.47:1314.16]) is less than half of R1 (52.49

[2.91:1728.93]), and in hard designs, R3 median TSubmit (23.55 [1.72:892.73]) is nearly half of R1 (42.54 [2.17:1728.5]). This highlights the impact of repetition on response time, showing participants' increased efficiency with task familiarity. This is consistent with findings in skill acquisition literature [30]. The observed result supports the notion that repeated exposure leads to improved performance and reduced cognitive effort [31].

Gender*:* A significant difference in TSubmit was observed among different designs based on gender. In easy design, where the highest TSubmit median is for Design III in both genders (male: 22[5-223] & female: 23[4-301]) when compared to Design I (male: 19[3-1465] & female: 20[7-665]) and Design II (male: 17[5-636] & female: 20[3-468]). It is clear from the interaction plot (see Fig.5) that males and females spent nearly the same amount of time in Design I and III, but males spent much lower TSubmit in Design II. The difference in TSubmit among genders is significant, although the interaction is not. Also, in medium and hard designs, the gender as well as the designs are significant as in medium designs the TSubmit in Design II (male: 65[3-592]& female: 72[5-589]) higher than other designs and in hard designs, Design I has the lowest TSubmit (male: 34[2-436]& female: 40[2-489]) while Design II has the highest TSubmit (male: 52[2-207]& female: 52[2-527]). The gender-based differences in response time highlight the nuanced interaction between design complexity and user characteristics (i.e., gender). Consistent with previous research on gender disparities in cognitive processing [32], e.g., information processing styles [33]. These findings highlight the importance of adopting inclusive design practices that accommodate diverse user demographics and cognitive preferences, aligning with the universal design principles [34]. Further research exploring the underlying factors driving these differences could provide deeper insights into designing more inclusive and user-centric visualizations.

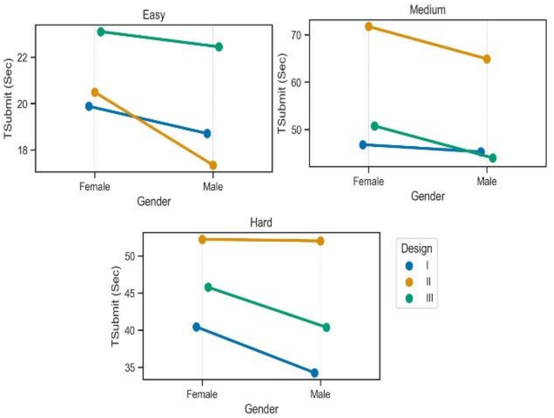

Figure 5. Interaction plot of TSubmit among different genders.

3.2 Interaction

Designs Complexity*:* In easy designs, there is no significant difference in Click Count among different designs. However, in medium designs, Click Count is significantly lower in Design I and Design III (2.79±2.19 and 2.82±2.34, respectively) than in Design II (2.97±2.86). Notably, Design I has the lowest maximum Click Count (2.0[2.0:23.0]), which is nearly half the maximum of Design II (2.0[2.0:55.0]). Furthermore, in hard designs, Design I has a significantly lower Click Count (2.87±2.34) when compared to Design II (3.01±2.64). In easy designs, users interact similarly. However, medium-level designs have notable differences. Design II has higher interaction and Click Count, making it more effective in engaging users. This highlights its potential to enhance user engagement and lerning efficiency. In easy designs, where the cognitive load is minimal, no significant disparity in Click Count among different designs is observed, aligning with prior studies emphasizing simplicity in user interactions [35]. However, significant differences emerge in medium and hard designs, with Design II exhibiting a higher Click Count than other designs. This finding resonates with literature highlighting the importance of interactive elements in engaging users and facilitating task learning [36].

Repetition: In easy designs, there is no significant difference in Click Count among different designs. However, in medium designs, Click Count is significantly lower in Design I and Design III (2.79±2.19 and 2.82±2.34, respectively) than in Design II (2.97±2.86). It is worth noting that Design I has the lowest maximum Click Count (2.0[2.0:23.0]), which is nearly half the maximum of Design II (2.0[2.0:55.0]). Furthermore, in hard designs, Design I has a significantly lower Click Count (2.87±2.34) when compared to Design II (3.01±2.64). Across different complexities, the impact of repetition on Click Count is observed, particularly in medium and hard designs. Consistent with theories of skill acquisition and automation [37], participants exhibit increased interaction efficiency with task familiarity, as shown by higher Click Count in Design II across all complexities. Thus, the findings underscore the importance of iterative design approaches that facilitate user learning and skill development over time.

Gender: It is clear that in all designs females have a significantly higher Click Count than males. The average Click Count in females in all designs 3±2 is significantly higher than in males 2±1. Moreover, there is a significant difference among designs where Design I has the lowest average Click Count (male: 2.6±2 & female: 3.2±3), whereas Design II has the highest average Click Count (male: 2.8±2 & female:

4.2±6). There is a significant difference between gender, where males have a significantly lower Click Count in all designs (Design I: 2.5±1, Design II: 2.8±2 and Design III: 2.7±2) than in females (Design I: 3.3±4, Design II: 3.4±4 and Design III: 3.3±3). The gender-based differences in Click Count reveal disparities in user interaction behavior. Females consistently exhibit higher Click Count across all designs, suggesting potential gender-specific differences in interaction patterns. These findings align with previous research highlighting variations in cognitive processing styles between genders [38]. Thus, the observed differences in Click Count among designs underscore the importance of considering gender-specific preferences and cognitive styles in visualization design [39].

3.3 Accuracy

Designs Complexity: The analysis of accuracy rates across different design complexities reveals significant differences in task performance. In easy designs, where the accuracy rates are consistent among all designs with around (6%) of wrong answers, no significant variation observed. However, substantial differences emerge in accuracy rates among different designs in both medium and hard complexity designs. In medium designs, Design I and Design III demonstrate significantly lower percentages of wrong answers (10.2% and 11.5%, respectively) than Design II (19.5%), indicating their effectiveness in facilitating accurate task learning.

Similarly, in hard designs, Design I exhibits the lowest percentage of wrong answers (15.4%) compared to Design II (17.3%) and Design III (21.1%), suggesting it outperforms task accuracy. Therefore, Design I is the preferred option across medium and hard complexities based on its consistently higher accuracy rates than the other designs. Examination of accuracy rates across varying design complexities sheds light on the nuanced relationship between design complexity and user performance. In easy designs, where the cognitive load is minimal, consistent accuracy rates are observed across all designs, aligning with prior research highlighting the simplicity of tasks in this category [40]. However, in both medium and hard complexity designs, significant differences in accuracy rates appeared, suggesting the impact of design on user performance in completing target tasks. These findings resonate with studies emphasizing the importance of design optimization in enhancing user performance and satisfaction [41]. This underscores the critical role of design considerations, particularly in complex tasks, where design choices significantly influence task accuracy and user experience.

Repetition: In easy designs, there is no significant difference in accuracy between all designs among all repetitions. In medium design, Design II has the lower accuracy in all repetitions as it has the highest percentage of wrong answers in R1 157(10.4%), which is more than twice that in Design I and Design III [68(4.5%) and 74(4.9%), respectively]. Concerning R2, Design I has the lowest percentage of wrong answers 34(2.3%), which is nearly half that in Design II and Design III [64(4.3%) and 58(3.9%), respectively]. Moreover, in R3, Design III has the lowest percentage of wrong answers 41(2.7%), when compared to Design I and Design II [51(3.4%) and 72(4.8%), respectively]. The difference in accuracy a cross medium designs is significant in all repetitions. In hard designs the accuracy of hard designs is nearly the same, and there is no significant difference in the percentage of wrong answers in R1 and R2 in all designs, around 7% and 5% respectively. However, in R3, the accuracy is significantly varied among different designs as Design III has the highest percentage of wrong answers 117(7.8%), which is one and half times Design II’s 81(5.4%) and more than twice in Design I, 53(3.5%). The impact of repetition on accuracy rates reveals interesting patterns in user performance across varying design complexities. In easy designs, where tasks are straightforward, no significant difference in accuracy is observed across repetitions, consistent with prior studies emphasizing performance stability in low-complexity tasks [42]. However, in medium designs, significant differences appear in accuracy rates across different repetitions, highlighting the role of task familiarity in influencing user performance. Design II exhibits lower accuracy rates across all repetitions, suggesting potential challenges in task completion. Comparing both results, it's evident that while design complexity plays a significant role in influencing accuracy rates, the impact of repetition is more pronounced in medium and hard complexity designs. This suggests that while design optimization is crucial, iterative refinement based on user feedback and task repetition is equally essential in enhancing user performance and satisfaction across different complexities.

3.4 Preference

Designs Complexity: The study found that Design II had the highest percentage of general preference and understandable design, with (41.8%) and (39.26%), respectively. This is nearly twice the percentage found in Design III, which had (21.48%) and (23.05%), respectively. Additionally, the choice of appeal in Design II and Design III was more than one and a half times that of Design I. Specifically, Design II and Design III had (39.06%) and (34.38%) appeal, respectively, while Design I had (24.41%). Furthermore, the resulting percentage of the most memorable design was Design III 250, which is almost twice as high as in Design II, with (48.83%) and (27.34%), respectively, as well as in Design I, with (20.31%). Lastly, more than half of the participants believed that Design III could trigger behavioural changes (52.34%), which is almost four times that of Design II (15.43%) and Design I (9.38%).

Regarding the factors affecting preferences, about half of the participants chose the colored bar with icons (44.53%), while nearly one-third chose the colored bar with text (39.26%). The lowest percentage of participants (14.84%), chose the icons bar with icons. Only seven participants had other opinions about the factors affecting their preference. Analysing preference across different design complexities uncovers significant user preferences and perceptions disparities. Design II ranked as the preferred choice regarding general preference, understandability, and appeal with substantially higher percentages than Designs I and III. At the same time, Design III is preferred for memorability and potential for triggering behavioral changes. These findings align with previous research highlighting the importance of user-centered design principles, where intuitive and visually appealing charts are favored by users [43]. The higher preference for Design II suggests its effectiveness in engaging users and promoting positive responses, resonating with studies emphasizing the significance of aesthetics and usability in visualization design [44].

Age: More than one and a half times group 45-64, which has a percentage of memorable for Design III 17(30.91%). Moreover, in understandable, Design II has the highest percentage in the lowest age group 54(46.55%). which is more than three times prefer than Design III, 18(15.52%) under the same age group but in age 25-34, Design I 97(39.11%) is preferred nearly more than one and half times Design III 59(23.79%). Moreover, Design I 24(43.64%) is also preferred also in the largest age group, nearly four times more than Design III 11(20%). There is a significant difference between factors of preference and age groups where all age groups preferred colored bars with icons; nearly half of each group except the age group 25-34 less than 40% chose this preference 97(39.11%), while this age group preferred more colored bars with text 108(43.55%) this percentage is nearly one and half times the percentage of same choice in the age group 35-44 (see Fig.6).

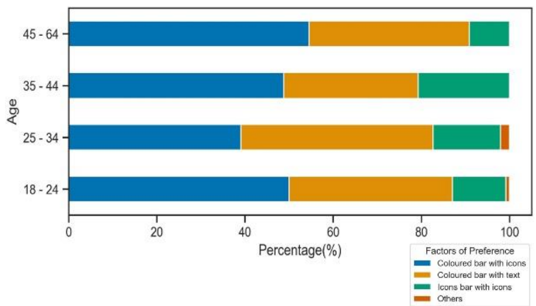

Figure 6. Percentage of factors of preference among different age groups.

Moreover, the percentage of those who have chosen icons bar with icons in the age group 35-44 is nearly twice 17(20.73%) that in the highest and lowest age group. The two intermediate age groups have a great percentage of choice icons bar with icons. The analysis of preferences across different age groups shows notable differences in design preferences and perceptions. Design II is favored among younger age groups, particularly regarding understanding ability. In comparison, Design I is preferred among older age groups, indicating potential age-related differences in cognitive processing and visualization preferences [45]. These findings align with theories of cognitive aging, where older adults may prefer simpler charts with clear navigation cues [46]. Additionally, the preference for colored bars with icons among all age groups underscores the importance of visual cues and intuitive design elements in enhancing user experience across diverse age demographics [47].

4. IMPLICATIONS

The study's findings underscore the importance of thoughtful design choices in visualizing correlational data. Tailored visualization designs can significantly impact task performance metrics such as accuracy and response time, suggesting potential improvements in data interpretation and analysis efficiency. Moreover, the influence of demographic factors like age on user preferences highlights the need for inclusive and user-centered visualization approaches. Specific design features, such as icons in colored bar charts, offer valuable insights for future research and practical applications in data visualization. Advancements in visualization techniques based on these findings can enhance decision-making processes in fields reliant on correlational data analysis. Ultimately, leveraging insights from this study can lead to improved data-driven decision-making and knowledge-sharing across various domains. By providing guidance on tailored visualization designs and considering demographic factors like age, our findings can enhance data interpretation and analysis.

5. CONCLUSION

This study examines the effectiveness of different colored bar chart designs in visualizing correlations between social media use, mental health, and family functioning. It emphasizes the importance of design choices in influencing response time, accuracy, user engagement, and preferences across various task complexities. The study highlights the influence of gender on user behavior and preferences, underscoring the need for personalized visualizations for diverse user demographics.

6. REFERENCES

[1] Alice, Plutino., L., Armellin., Andrea, Mazzoni., Roberta, Marcucci., Alessandro, Rizzi. (2023). Aging

variations in Ishihara test plates. Color Research and Application, doi: 10.1002/col.22877

[2] Alwuqaysi B, Abdul-Rahman A, Borgo R. The Impact of Social Media Use on Mental Health and Family Functioning Within Web-Based Communities in Saudi Arabia: Ethnographic Correlational Study. JMIR Form Res. 2024 Jan 16;8:e44923. doi: 10.2196/44923. PMID: 38227352; PMCID: PMC10828947.

[3] Andrew, S., McAvan., Andrew, S., McAvan., Yu, Karen, Du., Yu, Karen, Du., Alexis, Oyao., Alexis, Oyao., Stephanie, Doner., Stephanie, Doner.,Matthew, D., Grilli., Arne, D., Ekstrom., Arne, D., Ekstrom. (2021). Older Adults Show Reduced Spatial Precision but Preserved Strategy-Use During Spatial Navigation Involving Body-Based Cues.. Frontiers in Aging Neuroscience, doi: 10.3389/FNAGI.2021.640188

[4] Aniko, Simon., Mary, Guindy., Peter, A., Kara.,Tibor, Balogh., Laszlo, Szy. (2022). Through a different lens: the perceived quality of light field visualization assessed by test participants with imperfect visual acuity and color blindness. doi: 10.1117/12.2618638

[5] Ben, Steichen., Bo, Fu., Tho, Nguyen. (2020). Inferring Cognitive Style from Eye Gaze Behavior During Information Visualization Usage. doi: 10.1145/3340631.3394881

[6] Beverly Farrell (2022). The Development of an Instrument to Assess Strategical Information Processing Style. doi: 10.31390/gradschool_disstheses.338

[7] Cheng, Yang., Yiteng, Peng., Jing, Zeng. (2021). Research on cognition and application of icon complexity based on EEG. doi: 10.1007/S42486-021-00058-2

[8] Dragan, Rangelov., Hermann, J., Müller., Hermann, J., Müller., Michael, Zehetleitner. (2013). Visual search for feature singletons: multiple mechanisms produce sequence effects in visual search. Journal of Vision, doi: 10.1167/13.3.22

[9] Edy, Setiyo, Utomo. (2022). Analysis of the Mathematical Visualization Process of Female Students in Solving Contextual Problems Based on Cognitive Style. Phenomenon: Jurnal Pendidikan MIPA, doi: 10.21580/phen.2022.12.1.10850

[10] Eleanor, Eytam., Oded, Lowengart., Noam, Tractinsky. (2021). Effects of visual simplicity in product design and individual differences in preference of interactive products. Review of Managerial Science, doi: 10.1007/S11846-020-00391-3

[11] Epstein, N., Bishop, D., Levin, 1978. The McMaster model of family functioning. Publication Title:Journal of Marriage and Family Counseling Volume: 4

[12] G. A. Miller. The magical number seven plus or minus two: some limits on our capacity for processing information. Psychological Review, 63 2:81–97, 1956

[13] Ingrid, Maria. (2023). Universal Design and Inclusive Participation. doi: 10.4324/9781003266068-5

[14] Jianfeng, Wu., Dongfang, Jiao., Chunfu, Lu., Chengmin, Li., Xiaofang, Huang., Suzan, Weng. (2022). How Do Older Adults Process Icons in Visual Search Tasks? The Combined Effects of Icon Type and Cognitive Aging. International Journal of Environmental Research and Public Health, doi: 10.3390/ijerph19084525

[15] John, King., Kathy, Sonderer., Kevin, Lynch. (2016). Cognitive Benefits of a Simple Visual Metrics Architecture. doi: 10.1007/978-3-319-39396-4_29

[16] Julie, Baca., Daniel, W., Carruth., Alex, Calhoun., Michael, Stephens., Christopher, Lewis. (2021). Challenges in Evaluating Efficacy of Scientific Visualization for Usability and Aesthetics. doi: 10.1007/978-3-030-23535-2_29

[17] Kaifeng, Liu., Pengbo, Su., Hailiang, Wang., Da, Tao. (2022). Contextualizing Visualizations of Digital Health Information among Young and Older Adults Based on Eye-Tracking. Sustainability, doi: 10.3390/su142416506

[18] Karin, Eberhard. (2021). The effects of visualization on judgment and decision-making: a systematic literature review. doi: 10.1007/S11301-021-00235-8

[19] Leilani, Battle., Alvitta, Ottley. (2022). Testing theories of task in visual analytics. Interactions, doi: 10.1145/3529149

[20] Lichen, Zhang. (2023). Bar Charts, Butterfly Charts, Waterfall Charts, Dot Plots, Needle Plots, Area Bar Charts, Text Graphs, and Line Charts: Charts for Categorical Data. doi: 10.1007/978-1-4842-8609-8_4

[21] M. Elliott, C. Nothelfer, C. Xiong, and D. A. Szafir. A design space of vision science methods for visualization research. IEEE Trans. Vis. & Comp. Grap., 27:1117–1127, 2021

[22] Main, Uddin, Rony., Fan, Du., R., Rossi., Jane, Hoffswell., Niyati, Chhaya., Iftikhar, Ahamath, Burhanuddin., E., Koh. (2023). Augmenting Visualizations with Predictive and Investigative Insights to Facilitate Decision Making. doi: 10.1145/3543873.3587317

[23] Makoto Takeuchi; Kentaro Yumita; Rentaro Yoshioka (2023). UI Design Method for Visualizing Sensor Data to Enhance User Understanding. doi: 10.1109/iconat57137.2023.10080807

[24] Mariya, E., Manahova., Eelke, Spaak., Floris, P., de, Lange. (2020). Familiarity Increases Processing Speed in the Visual System. Journal of Cognitive Neuroscience, doi: 10.1162/JOCN_A_01507

[25] Mark, Andrew. (2022). Cognitive challenges in complex system design. doi: 10.21606/drs.2022.152

[26] Mark, Harrower., Cynthia, A., Brewer. (2003). ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. Cartographic Journal, doi: 10.1179/000870403235002042

[27] Mary, P., Czerwinski., Nancy, Lightfoot., Richard, M., Shiffrin. (1992). Automatization and training in visual search.. American Journal of Psychology, doi: 10.2307/1423030

[28] Mengting, Zhang., Almut, Hupbach. (2020). Repeated encoding fosters retention of perceptual detail in visual recognition memory. Learning & Memory, doi: 10.1101/LM.052209.120

[29] Nor Sheereen, Z., Rozumah, B., 2010. Using the 12-item general health questionnaire (GHQ-12) to assess the psychological health of malaysian college students. Global Journal of Health Science 2, 73–80. URL:http://www.ccsenet.org/journal/index.php/gjhs/article/view/3852, doi:10.5539/gjhs.v2n1p73.

[30] Nozipho, Matumbu. (2023). A User-Centred Approach to User Interface Languages and Icons: Co-evaluation and Co-creation of Accessible Digital Mobility Services. doi: 10.1007/978-3-031-26155-8_12

[31] Phil, Reed., Tegan, Fowkes., M., Khela. (2023). Reduction in Social Media Usage Produces Improvements in Physical Health and Wellbeing: An RCT. Journal of technology in behavioral science, doi: 10.1007/s41347-023-00304-7

[32] Pil, Brix, Purup., Pil, Brix, Purup., Steffen, E., Petersen. (2021). Characteristic traits of visualizations for decision-making in the early stages of building design. Journal of Building Performance Simulation, doi: 10.1080/19401493.2021.1961864

[33] Reshma Pise, Bharati Ainapure (2022). User Interface Design With Data Visualization Technique. 227-251. doi: 10.4018/978-1-7998-9121-5.ch011

[34] S. Palan and C. Schitter. Prolific.ac—a subject pool for online experiments. Journal of Behavioral and Experimental Finance, 17, 12 2017

[35] Saeed, Asgharzadeh. (2022). Visualization of Information and Its Application in Information Management. doi: 10.52547/ijamac.1.2.20

[36] Sara, Tandon., Alfie, Abdul-Rahman., Rita, Borgo. (2022). Measuring Effects of Spatial Visualization and Domain on Visualization Task Performance: A Comparative Study. IEEE Transactions on Visualization and Computer Graphics, doi: 10.1109/TVCG.2022.3209491

[37] Saratchandra, Chatterjee. (2022). Culture and perception. doi: 10.4324/9781003051497-8

[38] Shady, Afrin, Jeesan., Han-Seok, Seo. (2020). Color-Induced Aroma Illusion: Color Cues Can Modulate Consumer Perception, Acceptance, and Emotional Responses toward Cooked Rice.. Foods, doi: 10.3390/FOODS9121845

[39] So-Yeon, Yoon., Yun, Jung, Choi., Hyunjoo, Oh. (2015). User attributes in processing 3D VR-enabled showroom: Gender, visual cognitive styles, and the sense of presence. International Journal of Human-computer Studies \/ International Journal of Man-machine Studies, doi: 10.1016/J.IJHCS.2015.04.002

[40] Sukanya, Khanom., Anneli, Heimbürger., Tommi, Kärkkäinen. (2015). Can icons enhance requirements engineering work. Journal of Visual Languages and Computing, doi: 10.1016/J.JVLC.2014.12.011

[41] The data for this paper was generated using Qualtrics software, Version 2023 of Qualtrics. Copyright © 2023 Qualtrics. Qualtrics and all other Qualtrics product or service names are registered trademarks or trademarks of Qualtrics, Provo, UT, USA. https://www.qualtrics.com

[42] Tien-Lung, Sun. (2003). An icon-based data image construction method for production data visualization. Production Planning & Control, doi: 10.1080/0953728031000118787

[43] Ting Dai, Joshua M. Rosenberg, Michael Lawson (2022). Data Representations and Visualizations in Educational Research. doi: 10.4324/9781138609877-ree148-1

[44] Umair, Afzal., Arnaud, Prouzeau., Lee, Lawrence., Timothy, D., Dwyer., Saikiranrao, Bichinepally., Ariel, Liebman., Sarah, Goodwin. (2022). Investigating Cognitive Load in Energy Network Control Rooms: Recommendations for Future Designs. Frontiers in Psychology, doi: 10.3389/fpsyg.2022.812677

[45] V., Viqueira, Pérez., D., De, Fez, Saiz., F., M., Martínez, Verdú. (2010). Colour vision: theories and principles. doi: 10.1533/9780857090195.1.3

[46] Valentina, Bianco., Marika, Berchicci., Federico, Quinzi., Rinaldo, Livio, Perri., Donatella, Spinelli., F., Di, Russo. (2020). Females are more proactive, males are more reactive: neural basis of the gender-related speed/accuracy trade-off in visuo-motor tasks.. Brain Structure & Function, doi: 10.1007/S00429-019-01998-3

[47] Xuedi, Qin., Yuyu, Luo., Nan, Tang., Guoliang, Li. (2020). Making data visualization more efficient and effective: a survey. doi: 10.1007/S00778-019-005883