Visual Informatics 4 (2020) 1–10

Contents lists available at ScienceDirect

Visual Informatics

journal homepage: www.elsevier.com/locate/visinf

Evaluation on interactive visualization data with scatterplots

Quang Vinh Nguyen [a,b,*], Natalie Miller [c], David Arness [c], Weidong Huang [d], Mao Lin Huang [e], Simeon Simoff [a,b]

[a] MARCS Institute, Western Sydney University, Australia
[b] School of Computer, Data and Mathematical Sciences, Western Sydney University, Australia
[c] School of Psychology, Western Sydney University, Australia
[d] Faculty of Transdisciplinary Innovation, University of Technology, Sydney, Australia
[e] School of Software, Faculty of Engineering & IT, University of Technology, Sydney, Australia

ARTICLE INFO

Article history:
Received 2 July 2020
Received in revised form 17 September 2020
Accepted 20 September 2020
Available online 26 September 2020

Keywords:
Multivariate data visualization
Multidimensional data visualization
Scatterplots
Scatterplot matrix
Controlled study

ABSTRACT

Scatterplots and scatterplot matrix methods have been popularly used for showing statistical graphics and for exposing patterns in multivariate data. A recent technique, called Linkable Scatterplots, provides an interesting idea for interactive visual exploration which provides a set of necessary plot panels on demand together with interaction, linking and brushing. This article presents a controlled study with a mixed-model design to evaluate the effectiveness and user experience on the visual exploration when using a Sequential-Scatterplots who a single plot is shown at a time, Multiple-Scatterplots who number of plots can be specified and shown, and Simultaneous-Scatterplots who all plots are shown as a scatterplot matrix. Results from the study demonstrated higher accuracy using the Multiple-Scatterplots visualization, particularly in comparison with the Simultaneous-Scatterplots. While the time taken to complete tasks was longer in the Multiple-Scatterplots technique, compared with the simpler Sequential-Scatterplots, Multiple-Scatterplots is inherently more accurate. Moreover, the Multiple-Scatterplots technique is the most highly preferred and positively experienced technique in this study. Overall, results support the strength of Multiple-Scatterplots and highlight its potential as an effective data visualization technique for exploring multivariate data.

1. Introduction

Scatterplots have been recognized as one of the versatile, polymorphic and generally useful techniques for showing in pairwise axes the correlation and patterns of low-dimensional data (Friendly and Denis, 2005; Packham et al., 2005) as well as overview of a large amount of data (Sedlmair et al., 2013). Unfortunately, scatterplot techniques do not handle well data sets with a high number of dimensions or a large number of items. The overplotting and overlapping of data points may hinder the accurate extraction of information (Cleveland and McGill, 1984). The above issues were identified and enhanced by synergizing tasks, data, and designs to create an effective scatterplots design's frame work (Sarikaya et al., 2018).

A scatterplot matrix presents all pairwise scatterplots of attributes on a single matrix formation. Having all attributes mapped in this formation means it is easier to scan horizontally and vertically to assess differences in relationships between multiple variables and make comparisons between dimensions (Cleveland and McGill, 1984). However, repeated comparison pairs at the upper- and lower-halves, as well as blank diagonal panels, results in much redundancy within matrix displays. While the empty diagonal plots can be utilized to present histograms or other information (Cui et al., 2006), traditional scatterplot matrix techniques may not provide interactive functions, such as linking, brushing, and zoom-in view which involve highlighting or de-emphasizing data across multiple scatterplot display panels.

Nguyen et al. extended the possibilities of single scatterplot and scatterplot matrix techniques with *Linkable Scatterplots* technique (Nguyen et al., 2013, 2016b) which has been used on various data sets in genomic data analytics on various data sets in genomic (Nguyen et al., 2014, 2016a). It provides flexibility and control over the user interface and a number of plots and visual mapping so that it can be arranged in a way that best meets users' data exploration goals. Depending on the nature of the data being represented and the intentions of the user, this technique can provide more plot panels than the single scatterplot and reduce the number of unnecessary plots - a major problem in

* Correspondence to: School of Computer, Data and Mathematical Sciences Western Sydney University Parramatta, NSW, Australia.
E-mail addresses: Q.Nguyen@westernsydney.edu.au (Q.V. Nguyen), Natmill303@gmail.com (N. Miller), D.Arness@westernsydney.edu.au (D. Arness), Weidong.Huang@uts.edu.au (W. Huang), Mao.Huang@uts.edu.au (M.L. Huang), S.Simoff@westernsydney.edu.au (S. Simoff).

https://doi.org/10.1016/j.visinf.2020.09.004

the scatterplot matrix methods, one which may contribute to cognitive overload.

The interactive visualization mantra of overview first, then allowing filtering, and lastly providing details on demand (Shneiderman, 1996; Wang et al., 2015), which raises interesting research questions whether the flexibility in choosing number of scatterplots, visual mapping and interactive option would provide superior to the analytical performance in comparison with the traditional ones. The interaction could refine visualizations and aid deep visual exploration and information discovery in the scatterplot visualizations (Ware, 2004). However, there is a lack of an empirical study to evaluate the performance and the user experiences of the scatterplot methods.

This article contributes an empirical study on three types of scatterplots visualizations, called Sequential-Scatterplots, Simultaneous-Scatterplots and Multiple-Scatterplots. Sequential-Scatterplots method displays the multidimensional data with a sequence of individual scatterplots where only one plot is shown at a time. Simultaneous-Scatterplots presents all data variables concurrently with a number of scatterplots. In comparison with Simultaneous-Scatterplots, Multiple-Scatterplots has greater flexibility in choosing a number of scatterplots and visual interaction options. Fundamentally, this evaluation is predicated on the proposed advantages of scatterplots, and informed by inherent links between human cognition and data visualization. This study investigates whether Multiple-Scatterplots would result in more efficient (time taken to complete tasks) and more accurate exploration of multivariate data compared to traditional single and scatterplot matrix, in those who have been identified as having little experience interpreting scatterplots, and those with a moderate amount of experience interpreting scatterplots. More specifically, the study evaluates whether the flexibility in visual mappings, interaction and providing plot panels on-demand in Multiple-Scatterplots can contribute to the user performance. This study also examines user preferences and experiences of interacting with each of the scatterplot-related visualization techniques.

The following hypotheses have been evaluated in this study:

H1. Flexibility in choosing a number of visual plots and greater interaction ability in Multiple-Scatterplots may help participants perform visual search and comparison more efficiently (i.e. time and accuracy) during visualization.

H2. There is a performance difference between experienced and novice groups of participants.

H3. Multiple-Scatterplots may improve user experience and preference when using scatterplot visualization techniques.

2. Scatterplot visualizations

Scatterplots remain a powerful approach in their simplicity to visualize data with a low number of dimensions. They are more effective than landscape visualizations for both visual search (Tory et al., 2007) and visual memory (Tory et al., 2009), especially when studying the correlation between two variables. However, the effectiveness of these elements deteriorates when the number of dimension increases, as well as the number of data points. Scatterplots have been studied extensively for non-dimensional reduction (Baldridge, 2010) and dimensional reduction data (Sedlmair et al., 2013).

2.1. Scatterplot matrix

A scatterplot matrix enables the display of pair-wise variables of all attributes inside a matrix of plot panels. Some recent works, such as Scatterdice (Elmqvist et al., 2008), provide interaction and brushing to support the better exploration of the data. However, scatterplot matrix visualization does not readily present categorical values due to a single type mapping on the axes. This limits the effectiveness when visualizing large multidimensional multivariate data that usually contain different variable types.

There have been several variations on the scatterplot matrix as well as a combination with other visualization techniques. Scatterplot matrix may be integrated with parallel coordinate visualization to examine the correlations between dimensions and coordinates (Albuquerque et al., 2009; Heinrich et al., 2012). Scatterplot matrix can be generalized by integrating different visualization techniques for variable types, such as scatterplots for pairs of continuous variables, heat-maps for pairs of categorical variables and bar-charts for pairs of categorical and continuous variables (Emerson et al., 2013; Im et al., 2013). Although the generalized scatterplot matrix techniques are superior on certain tasks compared to scatterplot matrix (Im et al., 2013), they also follow the matrix concept to display all pair-wise of attributes concurrently. This could be problematic when visualizing big datasets with a high number of dimensions and data items. There is unused display space in the scatterplot matrix visualization, and the high number of display plots, but not all be useful, which makes each plot panel too small to effectively display a large amount of information.

2.2. Multiple scatterplots

Multiple-Scatterplots techniques, such as *Linkable Scatterplots* (Nguyen et al., 2016b), could provide flexibility to the analysts in choosing the number of plot panels, mapping of variables on the axes, as well as visual attributes. By showing selected variables concurrently, it is more effective to compare the correlation of variables within the limited space while reducing the unnecessary and crowded presentation of all information as a scatterplot matrix. This technique also provides several ways to interact with the information, including highlighting, filtering, selecting, linking and brushing, modifying the visual properties or showing regression lines.

For example, Fig. 1 shows a *Linkable Scatterplots*' visualization with six panels on Crime data that includes 90 counties in North Carolina, United States for the period from 1981 to 1987 (obtained from https://rdrr.io/rforge/Ecdat/man/Crime.html). The dataset has 24 attributes with mixed numeric and categorical values.

The mappings on the x-axis and y-axis on plot panels 1 to 6 at the current exploration stage are (*People Per Square Mile, Crimes Committed Per Person*), (*Probability Of Arrest, Crimes Committed Per Person*), (*Probability Of Arrest, Percentage of Young Males*), (*Tax Revenue Per Capita, Weekly Wage of Local Governments Employees*), (*Tax Revenue Per Capita, Weekly Wage of State Employees*), and (*Tax Revenue Per Capita, Weekly Wage of Federal Employees*) respectively. Colors are mapped to Region (red $\rightarrow$ central, blue $\rightarrow$ west and green $\rightarrow$ other), shapes are mapped to Year (1981 to 1987). Regression lines are turned on to assist the correlation between the variables. The current visualization only shows the data of three years, including 1981 (O shape), 1984 ($\lozenge$ shape) and 1987 ($+$ shape).

The top three plot panels of Fig. 1 indicate that the west region tends to have the lowest crimes committed per person in comparison with the central region and the other region. The Figure also shows the correlation that the more People Per Square Mile contributes the higher Crimes Committed Per Person (see Scatterplot 1), and the Percentage of Young Males may increase moderately the Crimes Committed Per Person (see plot panel 3).

The bottom three plot panels of Fig. 1 were used to analyze the correlation between Tax Revenue Per Capita and Weekly Wage of Employees of Local, State and Federal Governments. The figure

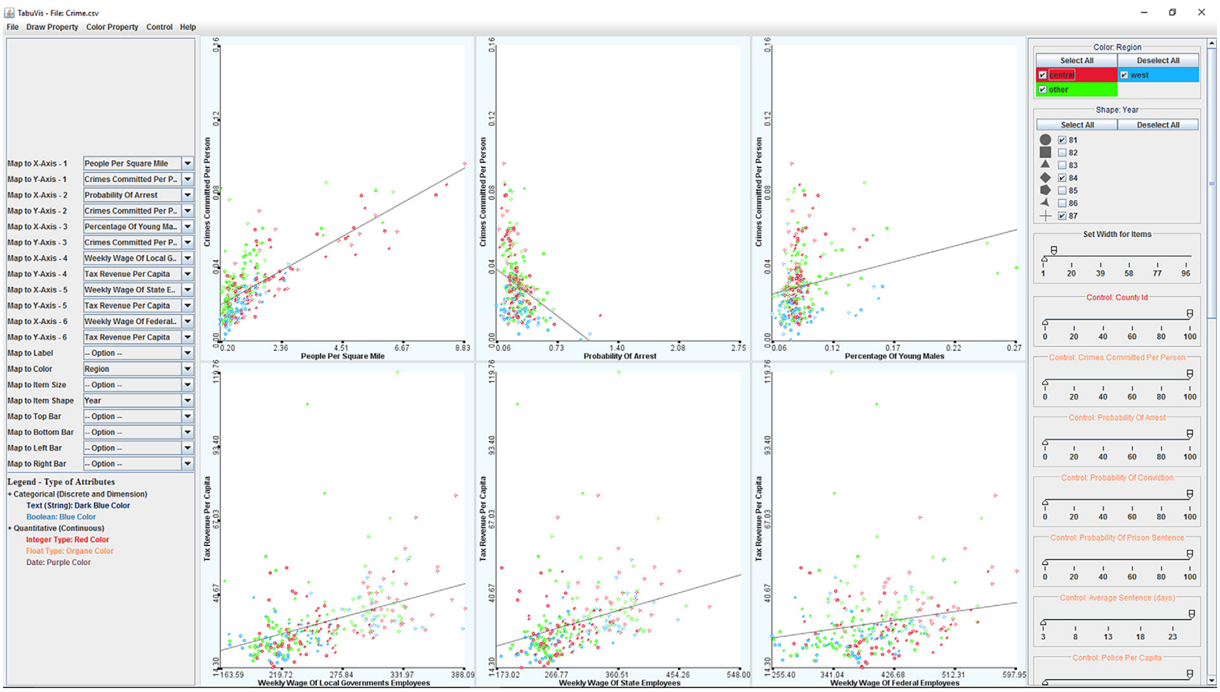

Fig. 1. An example of the linkable scatterplot visualization with six panels on the Crime in North Carolina over the period 1981–1987.

indicates the increase in wages in the years 1981, 1984 and 1987. There is no difference in wages for the three regions. It is interesting to note that the Tax Revenue Per Capita increases much faster when the Weekly Wage of Local and State Employees is higher (See Scatterplots 4 and 5), in comparison with the Weekly Wage of Federal Employees (see plot panel 6).

3. Controlled study

The value of a visualization technique should be evaluated by the degree to which it facilitates easy, efficient, accurate and meaningful interpretation of the story contained within the data (Few, 2013). This controlled study follows the most common way of data visualization usability which is evaluated through performance measures, particularly response time and accuracy (Huang et al., 2008, 2009). These measures are used because it is generally accepted that superior visualizations should facilitate faster and more accurate interpretation of the data (Zhu, 2007). Along with performance measures, the usability is also evaluated using subjective measures, particularly through finding out user's opinions about the utility and functionality of the data visualization technique, as well as their preference for using particular techniques (Huang et al., 2008; Zhu, 2007).

Our study did not use a single scatterplot because it was unable to display multivariate data. Similarly, a full classical scatterplot matrix was also not utilized for two reasons. Firstly, this visualization technique lent itself to overview but not a deep exploration of multivariate data (Nguyen et al., 2016b). Secondly, the scatterplot matrix technique would have limited the types of questions that could be asked as a result of its strength in displaying an overview of data rather than details, and this would have reduced the equivalence between conditions.

To ensure a fair comparison in this study, we adapted the single scatterplot and the scatterplot matrix with Sequential-Scatterplots and Simultaneous-Scatterplots respectively. The Sequential-Scatterplots technique showed interrelated simple scatterplots consecutively and might require participants to flick from one scatterplot to another in order to complete tasks. The Simultaneous-Scatterplots was an adapted version of the scatterplot matrix, which showed all scatterplots in a grid formation. We adopted *Linkable Scatterplots* (Nguyen et al., 2016b) to present Multiple-Scatterplots where the number of plot panels and axes and visual attributes can be defined by the users. Although multiple scatterplots could be generated with another tool, such as Tableau software (tableau.com), the *Linkable Scatterplots* was chosen in this study due to its greater capability in data interaction with mapping, linking, brushing and zooming among the plots.

The present study examined usability in the context of task accuracy and efficiency comparing Sequential-Scatterplots, Multiple-Scatterplots and Simultaneous-Scatterplots. We also examined the user's preferences and experiences of interacting with each of the scatterplot-related data visualization techniques. To ensure the consistency, the visualizations of Sequential-Scatterplots, Multiple-Scatterplots and Simultaneous-Scatterplots were generated using the same TabuVis tool (Nguyen et al., 2013).

3.1. Participants

A total of 40 participants, consisting of 31 females and 9 males completed the study, including 20 first-year and 20 fourth-year psychology students. Fresh first-year students did not usually gain experience with scatterplot and scatterplot matrix techniques and major statistical methods while fourth-year students likely learnt these techniques in their course study. The sample was chosen like this to form the two groups (Novice and Experienced), and group composition was confirmed via pre-test experience checklist. The Novice group consisted of 3 males and 17 females, with ages ranging from 18–45 years old ($M = 23.45$). The Experienced group consisted of 6 males and 14 females, with ages ranging from 20–43 years old ($M = 27.00$).

3.2. Design

A mixed-model design was adopted for this study, with two between-group factors (Novice and Experienced) and three within-groups, repeated-measures factors (Sequential, Simultaneous and Multiple-Scatterplots). The study measured time taken

Table 1
Descriptive statistics for accuracy and time.

Technique	Group	Accuracy		Time (s)	
		M	SD	M	SD
Sequential	Novice	.70	.17	41.36	15.32
	Experienced	.85	.13	31.69	7.89
	Total	.76	.16	36.52	12.98
Simultaneous	Novice	.53	.23	49.20	15.15
	Experienced	.60	.25	44.57	18.32
	Total	.57	.24	46.89	16.76
Multiple	Novice	.72	.22	46.36	15.79
	Experienced	.81	.15	36.14	8.89
	Total	.77	.19	41.25	13.67

to complete tasks (efficiency), accuracy on tasks (effectiveness), preference for each of the visualization techniques, and user-experience. Efficiency was measured as the total time taken to find the answer to all questions for each technique, and effectiveness was measured as the proportion of correct answer within the tasks for each technique. The groups, techniques and measured variables are shown in Table 1. In addition, a post-task user-experience (UEX) questionnaire invited participants to provide feedback about their experiences with each of the data visualization techniques, using 5-point Likert scales, as well as requesting open comments and a preference ranking for each of the data visualization techniques.

A multivariate data set chronicling crime data in New South Wales, Australia from 1995–2012 (in Excel Spreadsheets) was obtained from data.gov.au to be used as the basis for each task. The initial data set was large with almost 10,000 rows and contained many contingent variables (e.g. Years > Months, Offence Category > Subcategory). Given this complexity and a general lack of expertise in data analysis within the sample, the first step taken was to condense the data set by reducing the number of contingent variable categories for a more efficient visual representation to the participants. The above data processing was done with Microsoft Excel. Then, three separate data sets were created by systematically selecting unique year combinations from the larger data set within 15-year bands (e.g. 1995–2010, 1996–2011, 1997–2012). Each of these data sets was used randomly in one of three data visualization techniques. Separate data sets were required to ensure equivalence and minimize learning effects between the conditions.

Each data set was then represented using a corresponding visualization technique and five questions were formulated based on each. Tasks ranged in difficulty within each condition and were designed to be equivalent between conditions. Each task had three possible multiple choice answer options. This method was chosen to ensure participants with limited scatterplot experience could feel confident to provide an answer to the set tasks.

Pilot testing was conducted to ensure clarity and timing of the tasks. This resulted in minor word changes, such as changing the word "correlated" to "most similar" in task descriptions, to ensure tasks were understandable by participants with a diverse range of scatterplot experience. Pilot testing also resulted in minor changes to the colors and shapes utilized to ensure readability of the scatterplots.

3.3. Stimuli and measures

In addition to mapping to axes, color and shape were used to visually display variables across scatterplot panels. Interactive features such as filtering, width control functions and details on demand were emphasized to all participants as functions for exploring and navigating the data within the Multiple-Scatterplots technique. Effectively, participants in this condition had access to a series of related scatterplots and were required to complete a set of five tasks using three main functions as described — filtering, width control and details on demand. All participants started the task with all variables within the data set on display, and could then interact with the software functions to explore the data and answer the test questions.

Figs. 2 and 3 illustrate two examples of Multiple-Scatterplots visualization at a navigational stage. The images plot all offence categories in the Inner Sydney, Canterbury-Bankstown, Central Western Sydney, Central Coast, Mid North Coast, Hunter, Northern Beaches and Richmond area for the selected years. The sample questions included "Which offence category was the highest overall?", with a more challenging task question being "Overall, what happens to crime rates in Inner Sydney between 1995 and 2010?".

Fig. 4 shows an example of the visualizations of Sequential-Scatterplots in the study. The images plot all offence categories in the Eastern Suburbs, St George-Illawarra and the Central Coast in 2008 and 2011. Task questions relating to this set of scatterplots included, for example, "Which offence category increased the most between 1996 and 2011?", and a harder question was "Which offence category is most similar between 1996 and 2011?".

Fig. 5 shows an example of the visualization of Simultaneous-Scatterplots, which involves a series of four related scatterplots displayed in a grid formation. The task questions varied in difficulty, with some requiring participants to focus on one of the scatterplots in the grid, and others that required a comparison across scatterplots. For example, questions included "What year was arson and property damage highest in St George-Illawarra?", and a more challenging question was "Which year does the number of liquor and drug offences surpass the number of homicides and assaults in Inner Sydney?".

3.4. Post-test questionnaire

The post-test questionnaire consisted of four items related to aspects of user experience with the different scatterplot techniques, rated on a 5-point Likert scale (1 – Strongly Disagree, 5 – Strongly Agree and a neutral midpoint). The questions included "This scatterplot technique was easy to understand", "This scatterplot technique was easy to navigate", "This scatterplot technique was useful for completing the set tasks" and "I would choose to use this scatterplot technique with similar data sets in the future", which were adapted from other User-Experience surveys (Lund, 2001). There was also space at the end of each set of statements for participants to provide general feedback and comments about their experience using each of the techniques. Finally, participants were asked to rank each scatterplot technique in order of preference (1 = Most preferred, 3 = Least preferred).

3.5. Apparatus and materials

All participants worked on a HP ProBook laptop with a 15.6-inch matt screen display. All scatterplots used in this study were made using the same TabuVis tool (Nguyen et al., 2013), but the Sequential-Scatterplots and Simultaneous-Scatterplots were presented to the participants as images using Windows Photo Viewer. Participants were told they were able to flick between scatterplots in the Sequential-Scatterplots condition. Multiple-Scatterplots were run and displayed to participants using the tool to enable to mapping on multiple panels. Sample tasks, condition tasks and the post-study questionnaire were provided to participants on separate pieces of paper and participants were provided with a pen to record their answers. Time taken to complete each task was measured using a stopwatch and recorded on a separate document by the experimenter.

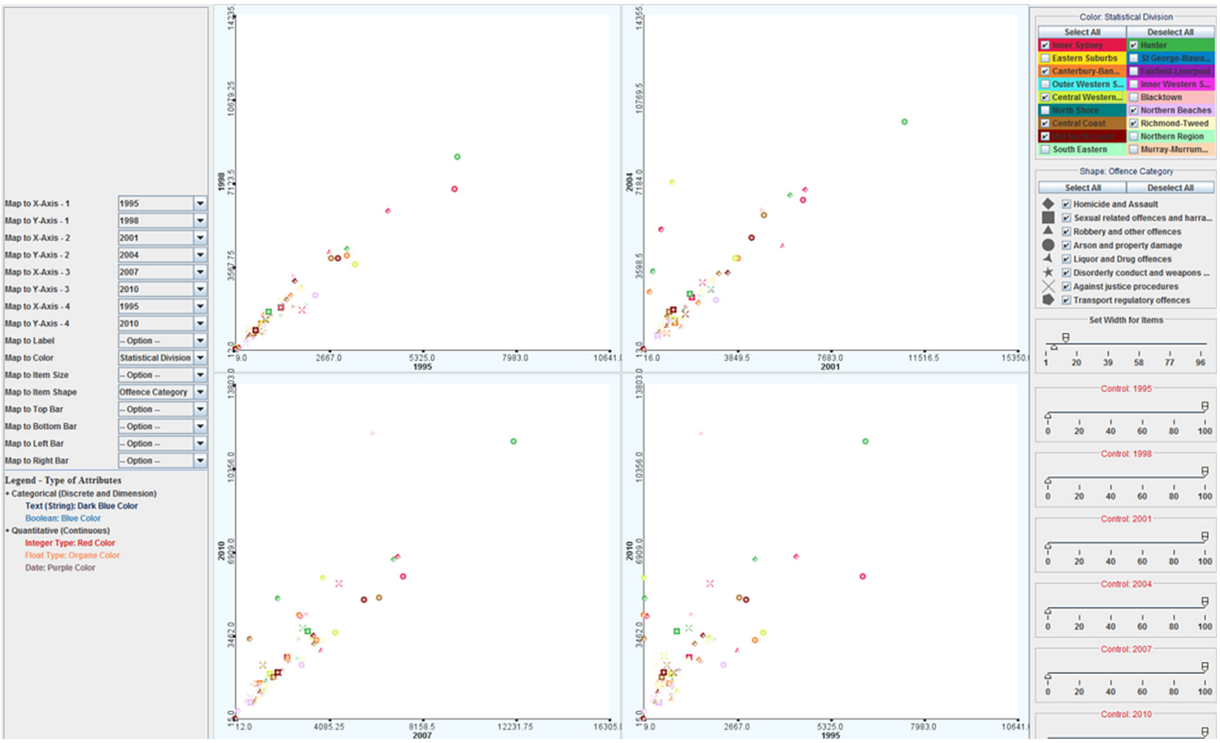

Fig. 2. An example of multiple scatterplots used in the study with 4 plot panels. The users can select mapping axes and visual properties to suit their exploration need, e.g. color to Statistical Division variable and shape to Category variable.

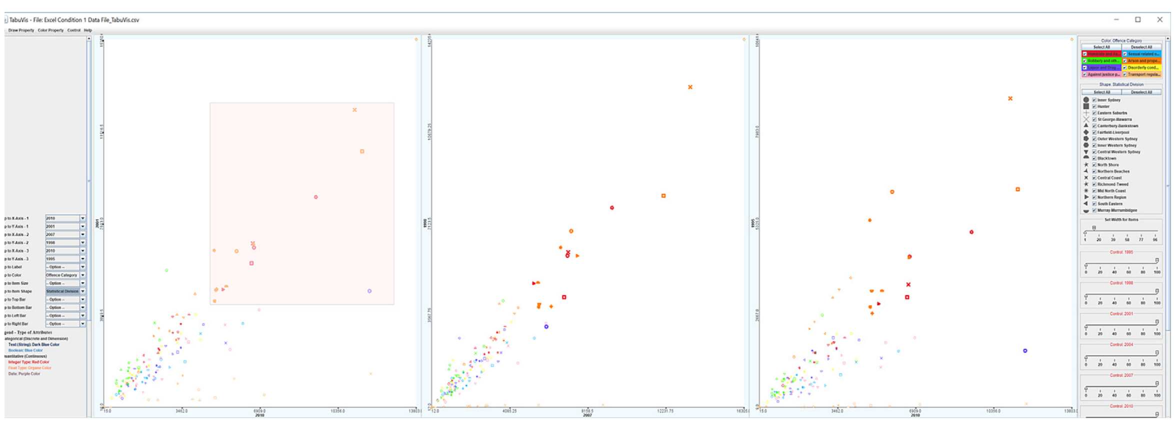

Fig. 3. Another example of the visualization with Multiple Scatterplots with 3 plot panels where the axes were assigned to different variables (e.g. Years). This example uses the same data set as in Fig. 2. The visualization shows different visual mapping color to Category variable and shape to Statistical Division variable. The user is selecting items for comparing across the plots.

3.6. Procedure

The within-subjects design meant that all participants completed the same conditions (Sequential-Scatterplots, Multiple-Scatterplots, and Simultaneous-Scatterplots), as well as the post-study questionnaire. All trials were undertaken in a quiet laboratory room on an individual basis with minimal distractions and controlled lighting conditions. The experimenter sat approximately one meter behind the participant whilst they were completing the tasks. The participant faced away from the experimenter, positioned approximately 40 cm from the laptop screen, in order to restrict any distractions or accidental facial recognition feedback from the experimenter.

Prior to the commencement of the first trial, all participants completed a 15-min training phase to introduce them to each of the visualization techniques in the study. The training phase involved being read standardized training instructions and being shown a demonstration of how to navigate each of the data visualization techniques. As well as explaining the static control techniques of Sequential-Scatterplots and Simultaneous-Scatterplots, the experimenter demonstrated how to map the data in the interactive Multiple-Scatterplot technique. Participants were shown an example data set, invited to explore the data and complete a set of sample questions using each of the scatterplot techniques. During this phase, participants were encouraged to ask any questions to clarify their understanding of using each of the data visualization techniques. Once participants felt they were confident using each of the techniques, the testing phase began.

In the testing phase, participants were required to complete a set of five tasks for each of the three scatterplots techniques. The researcher explained to participants that they would have a maximum of two minutes to complete each task. They were encouraged to complete tasks as accurately and quickly as possible,

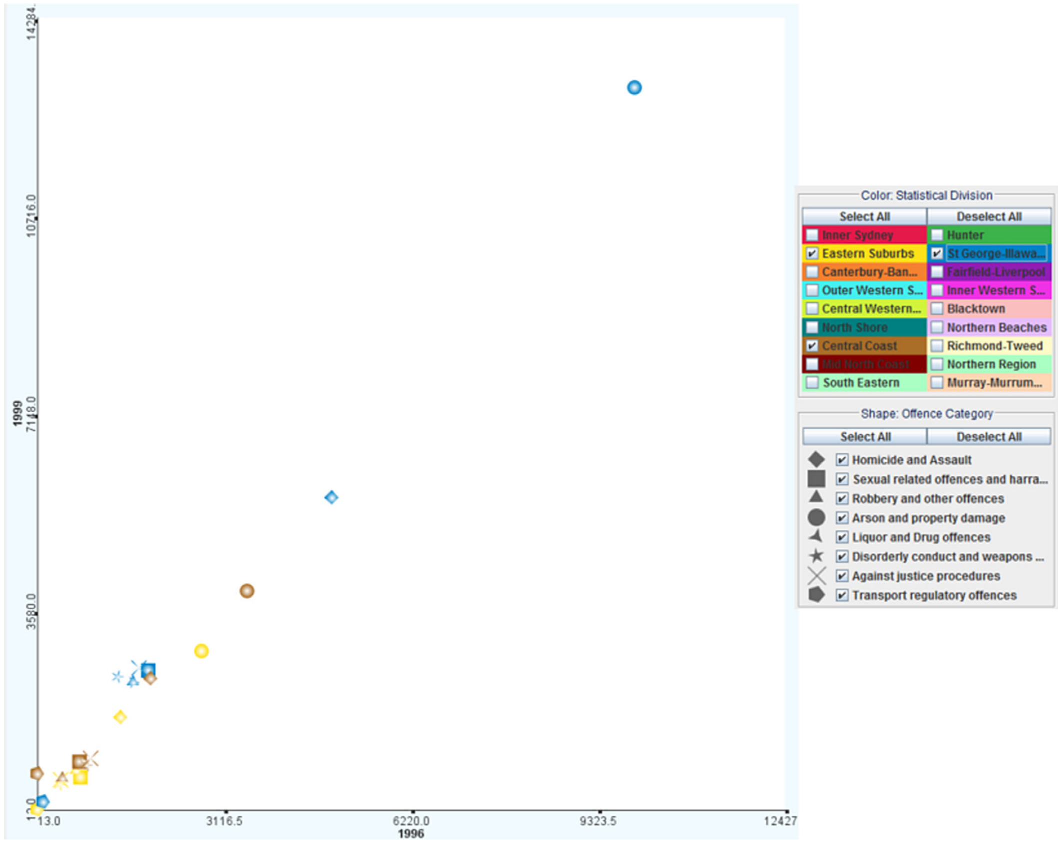

Fig. 4. An example of one scatterplot used in the Sequential Scatterplot experiments. In the experiment, these scatterplots were presented separately with participants free to switch between them.

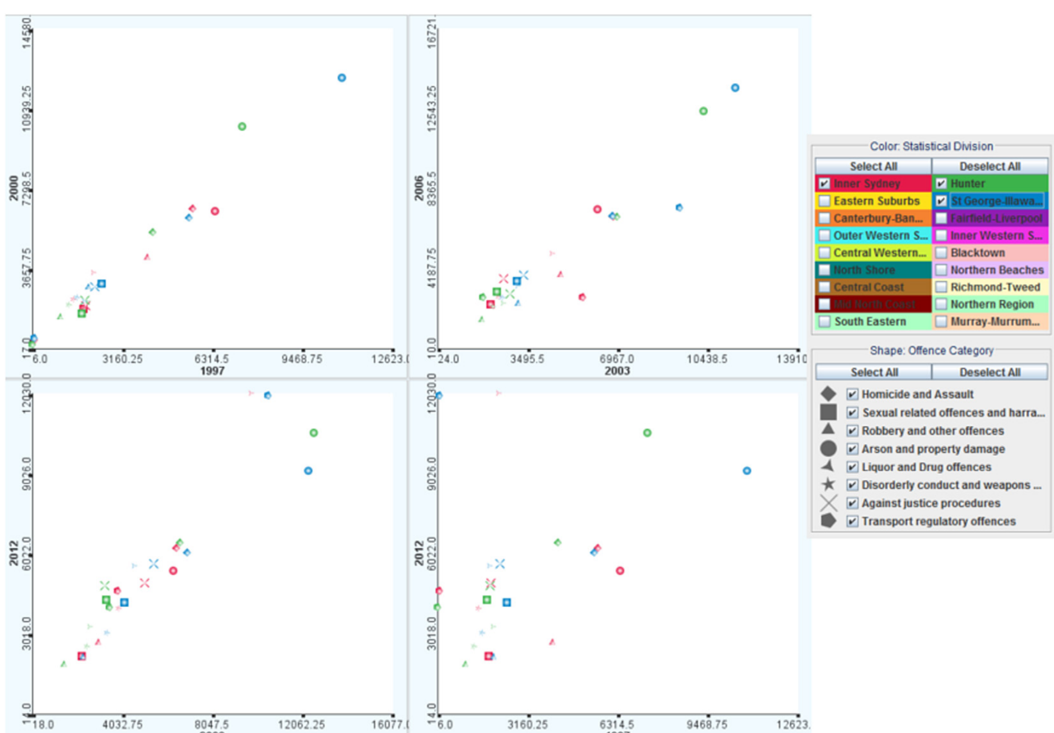

Fig. 5. An example of Simultaneous Scatterplots used in the study. The Figure shows values of all seven variables (years) in limited four scatterplots rather than a matrix of 7 × 7 scatterplots. This visualization does not provide the flexibility for the user to choose the number of plots, mapping axes and visual properties.

and to let the experimenter know when they were commencing reading of each task so the timer could be started. They were required to say 'Stop' when they were confident they had found the answer, but prior to recording it on the task sheet. The experimenter would record the time taken to complete the task and reset the timer before the participant began the next question. After all tasks had been attempted, participants were invited to complete a post-study questionnaire regarding their experiences and preference for using each of the techniques for exploring large data sets.

3.7. Results

3.7.1. Performance measures

A repeated measures mixed-model 2×3 multivariate analysis of variance (MANOVA) was conducted to determine if two dependent variables are associated with user performance (Time and Accuracy) differed significantly between novice and experienced users across the three data visualization technique conditions.

Initial data screening revealed no missing data and adequate sample size requirements. Non-significant Box's M, $p = .14$, indicated the assumption of homogeneity of covariance had been satisfied. The assumption of sphericity was satisfactorily met in accuracy, but violated in time, so a Greenhouse–Geisser correction was used for univariate tests. There were low to moderate correlations among the dependent variables, with no evidence of multicollinearity or singularity. While no multivariate outliers were identified as all Mahalanobis distance scores were below critical chi-square (χ^2) value of 22.46, $df = 6$, $p = .001$, two univariate outliers were identified in the time variable, with standardized residuals $\pm$ 3.29. However, the actual time taken by these participants was reasonably within expectations for the tasks.

3.7.2. Time and accuracy

Pillai's Trace was chosen as the most appropriate MANOVA statistic as it was considered to be powerful and generally more robust to violations of assumptions (Finch, 2005). Using Pillai's Trace criterion, there were significant main effects for both Group, $F(2, 37) = 5.77$, $p = .007$, $\eta\rho2 = .24$; and Technique, $F(4, 35) = 17.70$, $p < .001$, $\eta\rho2 = .67$. There was no significant interaction between Group and Technique, $F(4, 35) = .74$, $p = .571$, $\eta\rho2 = .08$. The initial MANOVA was conducted with outliers retained, and then re-run with outliers modified to be one unit more extreme than the next most extreme score (Tabachnick and Fidell, 2013). Pillai's Trace for data with transformed outliers maintained significant main effects of Group, $F(2, 37) = 5.65$, $p = .007$, $\eta\rho2 = .23$; and Technique, $F(6, 33) = 106.16$, $p = <.001$, $\eta\rho2 = .95$, and no significant interaction between Group and Technique, $F (6, 33) = 1.06$, $p = .41$, $\eta\rho2 = .16$. Results were equivalent, confirming the decision to retain outliers. Descriptive statistics for Time and Accuracy are presented in Table 1 and visually in Fig. 6, indicating greater accuracy and efficiency of experienced users in all visualization techniques which confirms Hypothesis 2 (H2). The figure also supports Hypothesis 1 (H1) that indicates the poorest performance in the Simultaneous-Scatterplots for both accuracy and time while Multiple-Scatterplots and Sequential-Scatterplots are quite comparable.

As a follow up to the significant MANOVA findings, separate one-way ANOVAs were conducted on Accuracy and Time, collapsed across experience levels and techniques, yielding four separate one-way ANOVA, with an adjusted alpha of.0125. Collapsed across technique, there was a statistically significant difference in accuracy $F(1, 38) = 7.74$, $p = .008$, $\eta\rho2 = .17$ but not in time taken on tasks, $F(1, 38) = 4.30$, $p = .045$, $\eta\rho2 = .10$. Collapsed across experience levels, one-way ANOVAs revealed that both accuracy, $F(2, 76) = 15.29$, $p < .001$, $\eta\rho2 = .29$; and time, $F(1.55, 58.80) = 15.71$, $p < .001$, $\eta\rho2 = .29$ were significantly different within the techniques.

Exploring these effects further, pairwise comparisons revealed significant differences in accuracy between Sequential-Scatterplots and Simultaneous-Scatterplots, MDiff = .210, $p < .001$, 95% CI [.104,.316], and Simultaneous-Scatterplots and Multiple-Scatterplots, MDiff = .20, $p < .001$, 95% CI [.08,.312], but no significant difference in accuracy between Sequential-Scatterplots and Multiple-Scatterplots. As displayed in Table 1, participants overall were less accurate using Simultaneous -Scatterplots than Sequential- and Multiple-Scatterplots.

Additional pairwise comparisons showed significant differences in time between Sequential-Scatterplots and Simultaneous-Scatterplots MDiff = −10.89, $p < .001$, 95% CI [5.81, 15.98]; and Sequential-Scatterplots and Multiple-Scatterplots, MDiff = −5.34, $p = .001$, 95% CI [1.94, 8.74]; but not between Simultaneous-Scatterplots and Multiple-Scatterplots. This demonstrates that participants were faster at completing tasks using Sequential-Scatterplots than the others. Taken all together, these results suggest that both *novice and experienced participants were most accurate on tasks using the Sequential-* Scatterplots *and Multiple-Scatterplots techniques, but fastest on tasks using the Sequential-Scatterplots* (H1).

Table 2
Descriptive statistics for preference ranking between groups and overall.

	Novice		Experienced		Overall	
	M	SD	M	SD	M	SD
Sequential	2.25	.72	2.55	.51	2.4	.63
Simultaneous	2.60	.50	2.45	.51	2.53	.51
Multiple	1.15	.37	1.00	.32	1.08	.27

Note: 1 = Most preferred, 3 = Least preferred.

3.7.3. Preference

Nonparametric tests were used to examine differences in preference ranking where data were ordinal and parametric assumptions were violated. A Kruskal–Wallis nonparametric test was conducted to determine if there was a significant difference in preference ranking for the three scatterplot techniques between novice and experienced participants. Results showed no significant difference across all groups in terms of ranking of Sequential-Scatterplots, χ^2 (1, N = 40) = 1.72, $p = .19$, $\eta\rho2 = .044$; Simultaneous-Scatterplots χ^2 (1, N = 40) = .88, $p = .35$, $\eta\rho2 = .02$; and Multiple-Scatterplots χ^2 (1, N = 40) = 3.16, $p = .08$, $\eta\rho2 = .08$. This suggests that *both novice and experienced participants ranked their preference for each technique similarly*.

As there was no statistically significant difference between how novice and experienced users ranked preferences for each technique, a Friedman nonparametric test was run to determine if preference ranking significantly differed between the three techniques overall. Results of this test showed that preference for the techniques significantly differed, χ^2 (2, N = 40) = 51.66, $p = .001$, $\eta\rho2 = 1.32$. As can be seen in Table 2 and Fig. 6, overall Multiple-Scatterplots were ranked most highly (1 = most preferred, 3 = least preferred) amongst all participants. *The scores suggest Multiple-Scatterplots technique was highly preferable in comparison with Sequential-Scatterplots and Simultaneous-Scatterplots techniques, which confirms the Hypothesis 3* (H3). Between the Sequential-Scatterplots and Simultaneous-Scatterplots, our study also indicates that experienced users have a slightly higher preference for Simultaneous-Scatterplots, whereas novice users have a slightly higher preference for Sequential-Scatterplots.

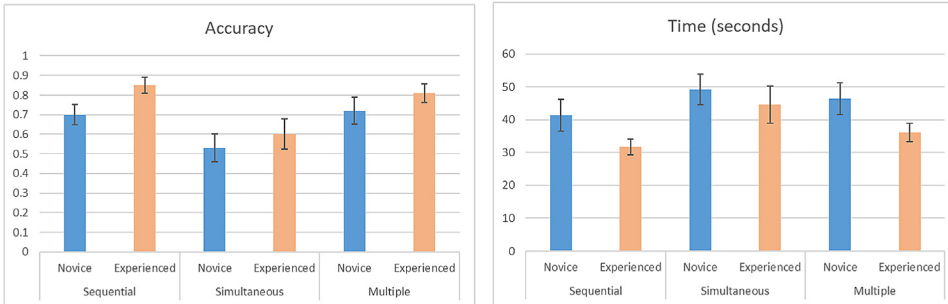

Fig. 6. Performance results with 95% confidence intervals: Accuracy (left) and completion time (right).

3.7.4. User- experience

Despite a small sample size for reliability analysis (N = 40), it was proposed that sample coefficients alphas could still be a robust estimate of population coefficient alphas with small sample sizes if the first factor of a Principal Components Analysis (PCA) loads highly (Cui et al., 2006). Assumptions relevant to all analyses (ANOVA, PCA, Internal Consistency) were checked prior to proceeding and were generally satisfactory. No missing data or univariate outliers were identified. No multivariate outliers were identified, as all Mahalanobis distance scores were below critical chi-square (χ^2) value of 32.91, df = 12, p = .001. The Shapiro–Wilk test indicated the normality assumption was met for both novice and experienced users in Sequential-Scatterplots and Simultaneous-Scatterplots User-Experiences, but not in Multiple-Scatterplots User-Experience. Violations to normality may be expected in Multiple-Scatterplots User-Experience, as histograms reveal that responses tended to cluster in the high range, possibly indicating a more positive user-experience using this technique.

A principal component analysis (PCA) with varimax rotation was conducted to examine factor loading of questionnaire item responses. The KMO Measure of Sampling Adequacy was sufficient at.70; and Bartlett's test of sphericity was significant (p < .001), confirming the factorability of the data and providing confidence to proceed. Results of this analysis show that the first factor in the PCA loaded at 4.10, which is typically considered quite a strong eigenvalue, indicative of highly consistent responses (Packham et al., 2005).

Results from the PCA lent confidence to run the planned Cronbach's alpha coefficients, as per Yurdugül (Yurdugül, 2008) to determine the internal consistency of responses. All Cronbach's alpha coefficients of the three Scatterplots techniques were above.85, indicating strong internal consistency, and validity of averaging questionnaire responses.

Results of the repeated measures ANOVA showed no significant main effect of group, F(3, 36) = .69, p = .56, $\eta\rho2$ = .054. This suggests that all users had a similar experience across each technique. However, there was a main effect of technique, F(2, 78) = 36.42, p < .001, $\eta\rho2$ = .48. Pairwise comparisons show that user-experience significantly differed between Multiple-Scatterplots and Sequential-Scatterplots, MDiff = .19, p < .001, 95% CI [.77, 1.62], and Multiple-Scatterplots and Simultaneous-Scatterplots, MDiff = 1.28, p < .001, 95% CI [.89, 1.67], but not between Sequential-Scatterplots and Simultaneous-Scatterplots.

Descriptive statistics displayed in Table 3 and Fig. 7, show that participants overall tended to *strongly agree with the user-experience items relating to Multiple-Scatterplots* (H3), indicating a highly positive user-experience, compared with essentially neutral experiences with the control techniques.

Table 3
Descriptive statistics for UEX responses relating to Scatterplots techniques.

Technique	M (SD)	95% CI
Sequential-Scatterplots UEX	3.32 (.92)	[3.02, 3.65]
Simultaneous-Scatterplots UEX	3.24 (.85)	[2.97, 3.51]
Multiple-Scatterplots UEX	4.51 (.67)	[4.30, 4.73]

Note: 1 = Strongly Disagree, 3 = Neutral, 5 = Strongly Agree.

3.7.5. Open-ended responses

Twenty-two participants provided comments in the open feedback section of the questionnaire. Responses received were collated on the basis of groups (novice and experienced), and techniques (Sequential-, Simultaneous-, Multiple-Scatterplots) with simple content analysis to identify themes and in particular similarities and differences in users' experiences.

Mirroring the indicators of user-experience and preference, most respondents expressed an inclination for the Multiple-Scatterplots technique. All respondents stated that they appreciated the interactive nature of Multiple-Scatterplots, which enabled control over navigation features. More specifically, nineteen respondents reported that the filtering function, which enabled them to hide irrelevant data and focus on important information, was the most useful feature for completing the tasks. Filtering also facilitated greater confidence in answers for a large portion of respondents. Increased confidence in correct answers using Multiple-Scatterplots was also evident in informal feedback from participants whilst completing the study. This aligned well with the statistical finding that the participants spent longer time to complete the tasks in Multiple-Scatterplots due to the exploration and learning processing, whilst they did the tasks more accurately thanks to the interaction capability. The flexibility of mapping and filtering in Multiple-Scatterplot technique also reflected the highest user preference and user experience in Fig. 6.

Feedback regarding the Simultaneous-Scatterplots technique indicated that many respondents found this technique to be problematic. Fourteen respondents reported that having all of the information available in four scatterplots was overwhelming and confusing to look at, and reported finding tasks more time-consuming. However, seven respondents in both groups reported finding this technique easier because of the ability to readily compare information across scatterplots, which meant they did not have to hold information in their memory. This could be due to the different levels of participants' capability in visual interpretation and working memory capacity when working with multiple views. The feedback explains the statistically low performance as well as low user preference and experience of this technique shown in Figs. 6 and 7 respectively.

Three respondents also mentioned that the Sequential-Scatterplots technique placed heavier demands on memory and attention, which made the tasks more difficult. Participants also

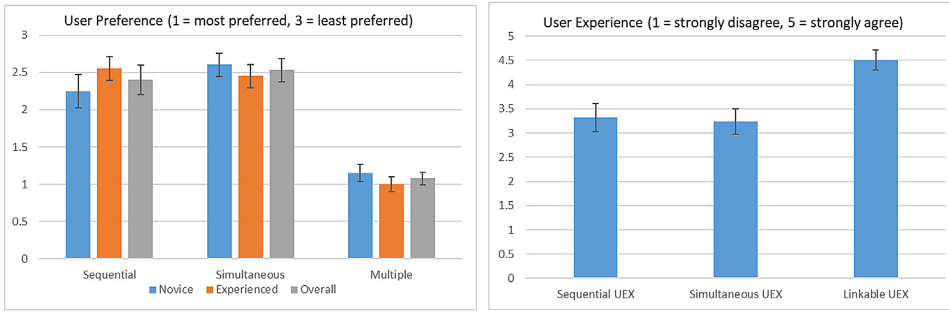

Fig. 7. User preference (left) and user experience (right) with 95% confidence intervals.

frequently made informal verbal comments about being unsure of their answers when using Sequential-Scatterplots and Simultaneous-Scatterplots techniques. Those comments are consistent with the statistical results that the participants spent less time to complete the tasks and achieved lower accuracy scores and user preference and experience when using the Sequential-Scatterplots.

4. Discussion

Conceptually, as an interactive technique, Multiple-Scatterplots had a range of advantages over Sequential-Scatterplots and Simultaneous-Scatterplots. Thus, it was anticipated that Multiple-Scatterplots would enable superior performance on tasks, resulting in higher accuracy, faster task completion time (efficiency), stronger preference, and more positive user-experience overall (H1–H3). For the most part, these expectations were supported in the present study, with the exception of task completion time. Whilst accuracy was highest using the Multiple-Scatterplots technique in comparison to control techniques, it did not result in the faster time taken to complete tasks. As expected, experienced participants overall were significantly more accurate using all scatterplot techniques than novice participants. However, in line with within-subject findings but contrary to expectations, there was no significant difference between novice and experienced participants in the time taken to complete tasks. Strong preference and positive user-experience using linkable scatterplots were supported in this study, highlighting the significant potential of the data visualization technique for use with large and multivariate data by a range of users.

In practice extracting information using data visualization techniques may involve a speed–accuracy trade-off, i.e. some techniques may require users to take more time in order to be most accurate (Lund, 2001). This could certainly account for the patterns of results found in the present study, with the newer and interactive technique Multiple-Scatterplots displaying high accuracy rates, but the more familiar and potentially simpler Sequential-Scatterplot technique, displaying both high accuracy and high efficiency.

There were two possible explanations for these findings. Firstly, due to the inherent simplicity of Sequential-Scatterplots, based on the single scatterplot, it was possible that the tasks in this condition were more straightforward than other conditions. Even though some participants suggested that looking across two scatterplots was difficult because it placed more demands on working memory, this also enabled unnecessary information to be excluded from view, essentially chunking vital information and resulting in less visual crowding for interpretation. Secondly, as a new and interactive technique, participants would have been least familiar with the Multiple-Scatterplots than the other two techniques. As a result, participants may have had to mentally work harder to remember functions that aid in the data exploration process, which may have made the Multiple-Scatterplots technique more time-consuming to navigate. Considering that there may be a speed–accuracy trade-off with Multiple-Scatterplots in the present study, the slower time taken to complete tasks does not necessarily detract from the overall usability. Further training using the Multiple Scatterplots format could potentially enable more efficient use of the technique.

Freedman and Shah's Construction-Integration model of graph reading suggested that graph reading experience and domain knowledge could affect how efficiently prior knowledge is activated about using graph formats, and by extension data visualizations, which could impact how efficiently information is extracted (Freedman and Shah, 2002). To this extent, it was expected that experienced participants would be more accurate and complete tasks faster than novice participants, as experienced participants may have been able to activate scatterplot schemata in order to complete tasks. This proposition was supported in the present study with results showing that experienced users were significantly more accurate across all techniques, and were faster at completing tasks than novice users. Putting the simple experience of interpreting scatterplots may have aided more accurate interpretation of scatterplots, but because the Multiple-Scatterplots technique was unfamiliar, both novice and experienced participants took more time to complete tasks using this technique.

Preference rankings and user-experience results demonstrated that participants had a stronger preference and more positive user-experience interacting with the Multiple-Scatterplots technique than the control techniques, which was in line with expectations. These measures provided valuable insight about how and why performance measures occurred and gave a strong indication of the perceived value of the techniques, which would not have emerged using performance measures alone (Huang et al., 2008). Even though Multiple-Scatterplots did not lead to the most optimal time performance, this technique did lead to higher accuracy. Feedback obtained through questionnaire data highlighted that higher accuracy was in a large part facilitated by the filtering function on Multiple-Scatterplots. As expected, filtering enabled users to remove unnecessary information from view in order to focus on aspects of the data pertinent to the tasks, which the majority of participants valued for exploring the data and completing tasks. It was also observed via informal feedback from participants whilst completing tasks that Multiple-Scatterplots aided confidence in response accuracy as a result of the filtering function. Whilst this was not directly measured in this work, it may be beneficial in future studies to include a confidence measure when comparing techniques.

Qualitative information obtained in questionnaires also highlighted key issues with each of the techniques. For example, a few participants noted that flicking from one scatterplot to another in the Sequential-Scatterplots conditions demanded more of their memory and attention, making the tasks more difficult. Similarly, the Simultaneous-Scatterplots technique was noted to be visually crowded with large amounts of data plotted, making the technique confusing for extracting information about specific data points, but useful for making comparisons across scatterplots. This provides further evidence of the need for intelligent, data visualization interfaces that can leverage perceptual strengths and compensate cognitive limitations, in order to aid effective and efficient data exploration and discovery.

Performance factors and subjective factors are often considered useful measures of the usability of a data visualization (Zhu, 2007). In the present study, usability was measured by examining performance factors, including the extent to which the Multiple-Scatterplot technique can facilitate easy, efficient, accurate and meaningful interpretation of the data (Few, 2013), as well as to how effectively this technique provides a positive and satisfying user-experience (Huang et al., 2008). In this sense, the present case study provides initial empirical support for the usability of the interactive Multiple-Scatterplot. The results of the present study indicate that the technique has significant potential as a user-friendly way that can facilitate effective data discovery and exploration with multidimensional data.

5. Conclusion

Multiple-Scatterplots technique hypothetically provides more plots for comparison than the single scatterplot while it also reduces the unnecessary number of plot panels of the scatterplot matrix. This technique also provides a rich interaction where the analysts can customize the visualization to suit their expectation and analytical requirements. This article presented a usability study of scatterplot methods, including Sequential-Scatterplots, Multiple-Scatterplots and Simultaneous-Scatterplots for explorative analysis of multidimensional data. Results of the controlled study indicated that Multiple-Scatterplots and Sequential-Scatterplots were more accurate than Simultaneous-Scatterplots while the completion time increased for Sequential- Scatterplots, Multiple- Scatterplots and Simultaneous-Scatterplots respectively. Multiple-Scatterplot gained significant potential as a user-friendly technique that could facilitate effective deep exploration with multidimensional data where completion time may not be crucial in the analysis process.

CRediT authorship contribution statement

Quang Vinh Nguyen: Software, Supervision, Writing - original draft. **Natalie Miller:** Formal analysis, Investigation, Writing - original draft. **David Arness:** Supervision, Validation, Writing - review & editing. **Weidong Huang:** Methodology, Writing - review & editing. **Mao Lin Huang:** Conceptualization, Writing - review & editing. **Simeon Simoff:** Conceptualization, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Albuquerque, G., Eisemann, M., Lehmann, D.J., Theisel, H., Magnorm, M., 2009. Quality-based visualization matrices. In: Presented at the Vision, Modeling, and Visualization (VMV).

Baldridge, G.R.R., 2010. The perception of correlation in scatter plot. Comput. Graph. Forum 29, 1203–1210.

Cleveland, W.S., McGill, R., 1984. The many faces of a scatterplot. J. Amer. Statist. Assoc. 79, 807–822.

Cui, Q., Ward, M.O., Rundensteiner, E.A., 2006. Enhancing scatterplot matrices for data with ordering or Spatial Attributes. In: Presented at the Visualization and Data Analysis.

Elmqvist, N., Dragicevic, P., Fekete, J.-D., 2008. Rolling the Dice: Multidimensional visual exploration using scatterplot matrix navigation. IEEE Trans. Vis. Comput. Graphics 14, 1141–1148.

Emerson, W., Green, W.A., Schloerke, B., Crowley, J., Cook, D., Hofmann, H., et al., 2013. The generalized pairs plot. J. Comput. Graph. Stat. 22, 79–91.

Few, S., 2013. Data visualization for human perception. In: The Encyclopedia of Human-Computer Interaction, second ed. ed: Interaction Design Foundation.

Finch, H., 2005. Comparison of the performance of nonparametric and parametric MANOVA test statistics when assumptions are violated. Methodology 1 (1), 27–38.

Freedman, E.G., Shah, P., 2002. Toward a model of knowledge-based graph comprehension. In: Proc. Diagrammatic Representation and Inference. In: ed: Lecture Notes in Computer Science, vol. 2317, pp. 18–30.

Friendly, M., Denis, D., 2005. The early Origins and development of the Scatterplot. J. Hist. Behav. Sci. 41, 103–130.

Heinrich, J., Stasko, J., Weiskopf, D., 2012. The parallel coordinates matrix. Comput. Graph. Forum 31, 37–41.

Huang, W., Eades, P., Hong, S.-H., 2008. Beyond time and error: a cognitive approach to the evaluation of graph drawings. In: Presented at the 2008 Workshop on beyond Time and Errors: Novel EvaLuation Methods for Information Visualization.

Huang, W., Eades, P., Hong, S.-H., 2009. Measuring effectiveness of graph visualizations: A cognitive load perspective. Inf. Vis. 8, 139–152.

Im, J.F., McGuffin, M.J., Leung, R., 2013. GPLOM: The generalized plot matrix for visualizing multidimensional multivariate data. IEEE Trans. Vis. Comput. Graphics 19, 2606–2614.

Lund, A., 2001. Measuring usability with the USE questionnaire12. Usability Interface 8, 3–6.

Nguyen, Q.V., Khalifa, N.H., Alzamora, P., Gleeson, A., Catchpoole, D., Kennedy, P.J., Simoff, S., 2016a. Visual analytics of complex genomics data to guide effective Treatment decisions. J. Imaging 2 (4), 1–17.

Nguyen, Q.V., Nelmes, G., Huang, M.L., Simoff, S., Catchpoole, D., 2014. Interactive Visualization for Patient-to-Patient Comparison. Genom. Inform. 12 (1), 263–276.

Nguyen, Q.V., Qian, Y., Huang, M.L., Zhang, J., 2013. TabuVis: A tool for visual Analytics Multidimensional Datasets. Sci. China Inf. Sci. 052105 (12).

Nguyen, Q.V., Simoff, S., Qian, Y., Huang, M.L., 2016b. Deep Exploration of multidimensional data with linkable scatterplots. In: proc. The 9th International Symposium on Visual Information Communication and Interaction. Dallas, Texas, pp. 43–50.

Packham, I.S.J., Rafiq, M.Y., Borthwick, M.F., Denham, S.L., 2005. Interactive visualisation for decision support and evaluation of robustness - In theory and in practice. Adv. Eng. Inform. 19, 263–280.

Sarikaya, A., Member, S., Gleicher, M., 2018. Scatterplots: Tasks, Data, and Designs. IEEE Trans. Vis. Comput. Graphics 24 (1), 402–412.

Sedlmair, M., Munzner, T., Tory, M., 2013. Empirical guidance on Scatterplot and Dimension reduction technique choices. IEEE Trans. Vis. Comput. Graphics 19, 2634–2643.

Shneiderman, B., 1996. The eyes have it: A task by data type taxonomy for information visualizations. In: Proc. the IEEE Symposium on Visual Languages.

Tabachnick, B.G., Fidell, L.S., 2013. Using Multivariate Statistics, 6th ed. Pearson Education, Boston.

Tory, M., Sprague, D., Wu, F., So, W.Y., Munzner, T., 2007. Spatialization design: Comparing points and landscape. IEEE Trans. Vis. Comput. Graphics 13, 1262–1269.

Tory, M., Swindells, C., Dreezer, R., 2009. Comparing dot and landscape spatialization for visual memory differences. IEEE Trans. Vis. Comput. Graphics 15, 1033–1039.

Wang, L., Wang, G., Alexander, C.A., 2015. Big data and visualization: methods, challenges and technology progress. Digit. Technol. 1, 33–38.

Ware, C., 2004. Information Visualization: Perception for Design, second ed. Morgan Kaufmann, San Francisco, CA, USA.

Yurdugül, H., 2008. Minimum sample size for Cronbach's coefficient alpha: a Monte-Carlo study. Hacet. Univ. J. Educ. 35, 397–405.

Zhu, Y., 2007. Measuring effective data visualization. In: Presented at the Advances in Visual Computing (ISVC 2007).