An Evaluation of Perceptually Complementary Views for Multivariate Data

Chunlei Chang[*] Monash University
Tim Dwyer [†] Monash University
Kim Marriott[‡] Monash University

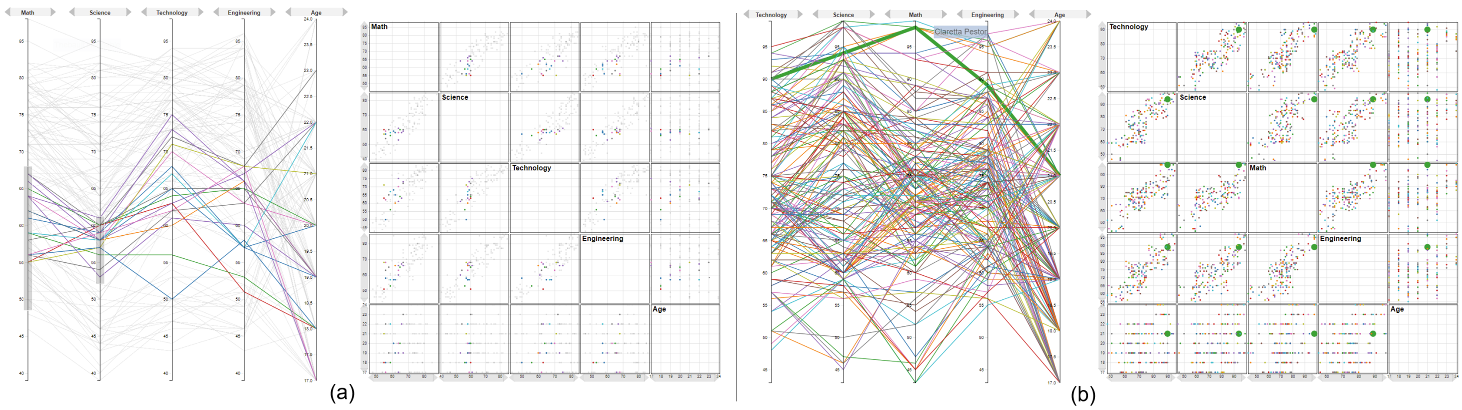

Figure 1: Samples of stimuli used in our Study. Side-by-side *PCP* and *SPLOM* with two difficulty levels: easy (a) and difficult (b). A multiple-selection filtering is applied on (a). Multiple *SPLOM* points and *PCP* lines corresponding to a single data item are highlighted in (b); Here, participants were required to find the candidate with the best average mark in over 100-200 candidates.

ABSTRACT

We evaluate the relative merits of three techniques for visualising multivariate data: parallel coordinates; scatterplot matrix; and a side-by-side, coordinated combination of these views. In particular we report on: (1) the most effective visual encoding of multivariate data for each of the six common tasks considered; (2) common strategies that our participants used when the two views were combined based on eye-tracking data analysis; (3) the finding that these views are perceptually complementary in the sense that they both show the same information, but with different and complementary support for different types of analysis. For the combined view, our studies show that there is a perceptually complementary effect in terms of significantly improved accuracy for certain tasks, but that there is a small cost in terms of slightly longer completion time than the faster of the two techniques alone. Eye-movement data shows that for many tasks participants were able to swiftly switch their strategies after trying both in the training phase.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Evaluation/Methodology

1 INTRODUCTION

Visualising high-dimensional data is a significant challenge for data visualisation designers and researchers as it tests the limits of human perception and cognition. Yet, with increasingly automated collection of data in almost every domain and hence ever more complex high-dimensional data becoming available, so too, effective visualisation and analysis tools become ever more important [23]. A single traditional visualisation idiom such as the humble scatterplot can effectively convey two data dimensions spatially. Colour or other visual channels can extend this to three or more data dimensions, though with less saliency than spatial mappings [8]. Beyond a certain degree of dimensionality (perhaps tens of dimensions) it may be essential to automatically select an interesting subset of dimensions, e.g. projection pursuit [16] or to rely on algorithmic techniques to reduce the dimensionality of the data, i.e. multidimensional scaling [4]. However, multidimensional scaling necessitates a loss of information in the projected view and the human analyst loses their direct connection to the underlying data.

There are, however, visualisation techniques that can map several (more than three but probably less than ten) data dimensions spatially without loss of information. Probably the most widely used of these techniques are scatterplot matrices (*SPLOM*s) [7] and Parallel Coordinate Plots (*PCP*s) [17].

Traditionally, scatterplots render two data-dimensions to orthogonal x and y axes. *SPLOM*s are a fairly logical extension of the traditional scatterplot to n data dimensions, by using small multiples of scatterplots to represent the full set of $n \times n$ pairs of data dimensions. *PCP*s use n axes arranged in a parallel sequence, using a polyline for each data point such that the data value in each dimension is indicated by the point at which the line crosses the corresponding axis. These two alternative techniques for multivariate data visualisation each have their own advantages and disadvantages, as explored in various studies described in detail in Section 2. Some of the more obvious problems for each technique can be summarised as follows:

SPLOM

- Scatter plot matrices require screen area proportional to the square of the number of dimensions.
- A single data element is represented by $n \times n$ points (one in each plot). Finding the points corresponding to a given data element necessitates interactive techniques such as brushing.

PCP

- Data distributions, correlations or other differences and similarities can only be easily compared on adjacent sets of axes.
- The lines connecting data points across axes can lead to significant clutter, especially when data dimensions are not well correlated causing the lines to cross in many places.

[*]e-mail: chunlei.chang@monash.edu
[†]e-mail: tim.dwyer@monash.edu
[‡]e-mail: kim.marriott@monash.edu

- While scatterplots an idiom encountered by most people in grade-school math classes, *PCP* is arguably less familiar to most people.

Which of the two techniques is better depends greatly on the particular analysis task. To overcome the disadvantages of the individual techniques, various hybrid designs have been proposed to combine the two different techniques into a single display. Typically, this requires quite sophisticated visual and interaction design that requires further training for users to use and interpret effectively. This paper evaluates a different, and arguably more straightforward approach to combining *SPLOM* and *PCP* visualisations into a single display that—to the authors' knowledge—has not been evaluated before. That is, simply placing *PCP* and *SPLOM* views side by side. Our study seeks to answer the question of whether users are able to easily switch between the two views in order to use the one most suited to the analysis task at hand. We do this by evaluating three conditions: parallel coordinates view only; scatterplot view only; and the two visualisations side-by-side. Our hypothesis is that speed and accuracy of participants completing different tasks using the combined view will be similar to the speed and accuracy of the better performing individual technique for the given task.

Our work in this area is motivated by a recent study evaluating the notion of *perceptually complementary* network visualisation by Chang *et al.* [5]. That work compared individual matrix and nodelink visualisations of networks against their side-by-side combination. Despite the two views being redundant in terms of both presenting the same information they found clear benefits to the combined view, as we discuss further in Section 2.

We ascribe this effect to the chosen pairings of network visualisations each supporting visual comprehension of different aspects of the data despite the two views being redundant in the sense that each on its own was sufficient—though perhaps not optimal—to complete the task.

Perceptually complementary visualisations are differentiated from a more common class of multiview displays that seek to show different subsets of data properties or data elements in each view. This latter class are called *informationally complementary* since each shows a different subset of the available information.

In this paper, we test if this principle of perceptual complementarity carries over to combined *SPLOM* and *PCP* views. In investigating perceptually complementary views on multivariate data we try to understand the following questions:

Q1: Which tasks are better supported by *PCP* or *SPLOM*?

Q2: Are there any advantages for *Combined* views when solving complex tasks?

Q3: Are people aware of the merits of different views? If so, what are their strategies to solve tasks.

The work presented in this paper may be beneficial to users in two distinct ways: first, our work shows that there is a perceptually complementary effect in terms of improved accuracy for some tasks; and second, eye-tracking data shows that in the combined condition, for almost all tasks, participants chose the more effective of the two views even without realising that they were doing so. In summary, our main contribution is an interactive *Combined* approach to visualise multivariate data with a complementary effect. The detailed contributions are that we:

- present a combined *PCP* and *SPLOM* view with novel interactions such as coordinated dragging of *PCP* axes and *SPLOM* cells;
- conduct a controlled study for six multivariate data analysis tasks;
- categorise the identified use strategies based on eye-tracking data;
- analyse the performance of *PCP* and *SPLOM* for six identified tasks;
- show how the six common tasks are supported by the effect of complementariness in the *Combined* view and what are the popular strategies used by participants.

The paper is organised as follows: after reviewing related work in Section 2, we present our design in Section 3. The user study is presented in Section 4. Finally, we analysed and discuss further results in Section 5 and draw conclusions in Section 6.

2 Related work

In this brief review, we focus on three subjects: previous work in the area of coordinated and complementary views for multivariate data visualisation (Section 2.1); empirical studies examining the visualisation of multivariate data (Section 2.2); and previous work that tries to combine aspects of *PCP* and *SPLOM* into hybrid visualisation systems (Section 2.3).

2.1 Coordinated and Complementary views

Multiple view systems with coordinated interactions allowing users to explore various aspects of a dataset have ascended from a mainstay topic in information visualisation research (e.g. [29]), to fairly standard practice in commercially available tools. For example, popular software tools like Tableau and Microsoft PowerBI allow people to create sophisticated tiled display dashboards, to filter the data interactively and view different attributes of the data using different standard visualisation techniques.

It would be fair to say that standard practice when creating multiview data displays is to choose sets of displays that are *informationally complementary*, in the sense that they show different sets of data items (such as an aggregated overview and a view of some subset of data showing more detail) or different sets of attributes (such as the side-by-side displays of timing and accuracy data that we have used to present our experimental results in this paper).

A less standard approach—and one argued by Chang *et al.* [5] to be undervalued in some situations—is to choose pairings of views that show the same data and the same attributes but in perceptually complementary ways. In their studies, Chang *et al.* tested analysis tasks for static and dynamic network (event sequence) data. Their paired views were an adjacency matrix and a node-link diagram for static data, and tiled matrix and Sankey diagrams for the event sequence data. Both views showed the full set of weighted edges, either through glyphs in matrix cells or link lines of varying thickness. However, as earlier network visualisation studies have indicated, these two views are markedly better for different types of tasks. That is, the matrix view is better when a precise comparison of link weight is necessary, and the node-link view is better for path following tasks [13]. Experimental results showed that most participants were able to use the two displays together in an effective way, such that the results of the combined views were at least as accurate as the best individual view for each task, and performance was only slightly slower.

This result is not completely unexpected, in fact it is predicted by results in perceptual psychology. For example, the so called "representational effect" described by Zhang and Norman [39] says that the same information shown in different ways may lead to different understanding. Within a single visualisation, it is also well known from visualisation research that redundant encodings within a single visualisation may improve readability [26]. However, to our knowledge, the Chang *et al.* study of network data visualisation was the first time that entire isomorphic displays (complementary visualisations) have been studied explicitly in the context of information visualisation. It raises two important questions prompting the study presented in this paper. First, can these results for network data visualisation be generalised to other data visualisation pairings? Second, the information visualisation research community has a long tradition of developing complex hybrid visualisation techniques such as those described for multivariate data in Section 2.3, but do we need to pay more attention to simple coordinated pairings?

2.2 Empirical evaluation of *PCP* and *SPLOM*

There have been many studies evaluating *PCP* and *SPLOM* techniques in isolation, but few that compare them directly, and none we are aware of that evaluate a side-by-side combination of the two. Johansson and Forsell [18] made a recent survey of evaluations of *PCP* and related techniques. Of the many studies surveyed, most relevant to this paper is one by Li *et al.* [22] comparing single scatterplots with *PCP* displays of 2 axes (only). In particular, they found that participants were better at accurately assessing degree of correlation between pairs of data dimensions with the scatterplot than with the *PCP*. The displayed stimuli were not interactive and the scatterplots and *PCP* were shown in isolation (a combined view was not tested).

Another study by Kanjanabose *et al.* [19] compares user performance, in terms of accuracy and response time, in the context of four different visualization tasks, using either *PCP* or *SPLOM* (but, again, not both together). Their results suggest that *PCP* is better than *SPLOM* for cluster, outlier and change detection.

A radial version of *PCP* called Stardinates was empirically tested against regular PCPs by Lanzenberger [21] - including a combination. The two views are very similar to one another (more similar than *PCP* and *SPLOM*), and hence, not particularly complementary. Thus, the result that few differences were found was not surprising although some benefit was claimed for the combined view.

2.3 Hybrid Scatterplot and Parallel Coordinates Visualisations

Many have identified similarities and complementary aspects of *PCP* and *SPLOM* displays of multivariate data. The visual saliency of scatter plots has long been considered by statisticians and others interested in data analysis. A classic paper by Tukey and Tukey [33] introduced metrics that they called *scagnostics*, for identifying notable structures in scatterplots. Wilkinson *et al.* [36] used scagnostics to analyse and order *SPLOM*s. Dasgupta et al. [10] extended this notion to *PCP* naming the equivalent metrics *pargnostics*. They also introduced the presentation of pairs of parallel coordinates axes in a matrix. Heinrich *et al.* [15] noted that such a complete PCP matrix, must introduce discontinuities in the line segments. So they introduced an alternate Parallel Coordinates Matrix (PCM) that presents a linear sequence of all pairs of axes. No controlled study was performed.

Siirtola [31] combined parallel coordinates with a tabular view to provide an informationally complementary overview and detail display. Note that the tabular view was not a scatterplot, rather each column was used to display individual attribute values for each underlying data item. The two views were coordinated such that interactions affected both views through selection brushing or reordering of dimensions. They performed an experiment comparing task effectiveness with and without linked interactions (both views were displayed at all times).

Yuan *et al.* [38] propose a parallel coordinates design which allowed users to interactively introduce points between two *PCP* axes at horizontal positions according to a third data dimension. Claessen and van Wijk [6] created an interactive tool for creating hybrid scatterplot and parallel coordinate displays, such that the axes of scatterplots could be extended in either direction into linked parallel coordinates axes. Cordeil *et al.* [9] have recently extended this concept to 3D in VR. Engelke et al. [12] presented a study on the visual assessment of relative data point distance in *PCP* and *SPLOM* in Catesian coordinate systems. Viau *et al.* [34] presented three novel and sophisticated approaches for achieving a tighter integration of *PCP* and *SPLOM* views through hybrid techniques for multidimensional visualization, sophisticated selection and morphing layout. While each of the above papers consider case studies, none of these hybrid systems have received controlled empirical user studies.

3 Visual and Interaction Design

For our study (described in Section 4) we are interested in measuring the effectiveness of *PCP*, *SPLOM*, and a *Combined* side-by-side representation of both *PCP* and *SPLOM* for multivariate data. For canonicity we used fairly standard visual designs for each of the techniques, as follows:

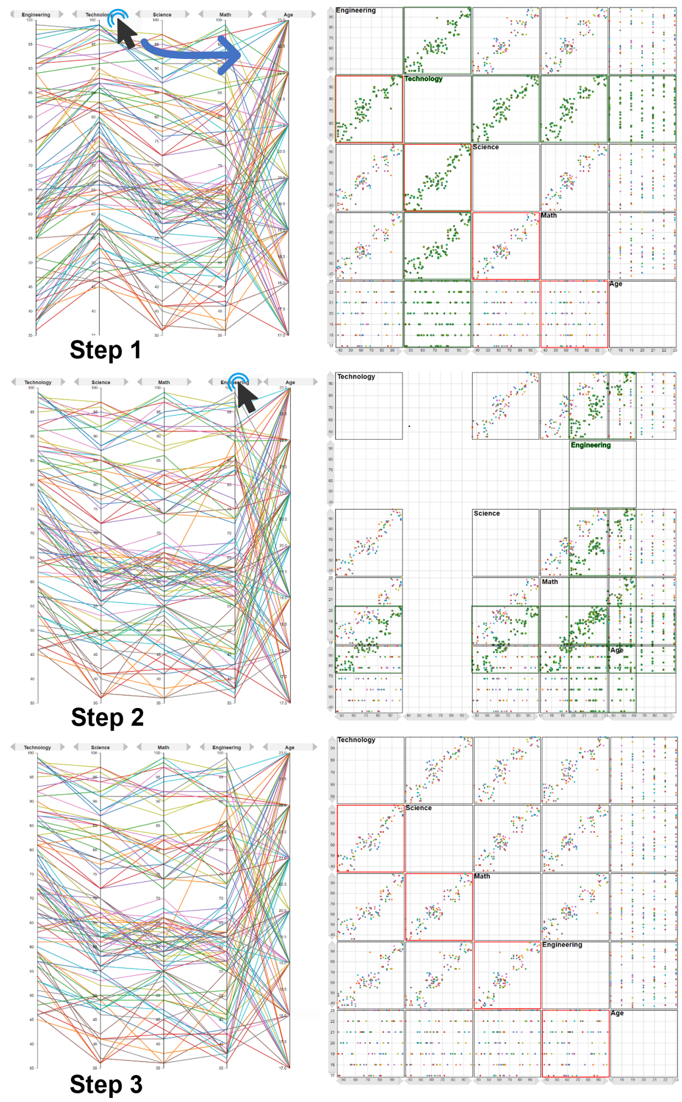

Figure 2: Re-ordering axes in *Combined* view. *Step 1:* User initiate dragging by clicking top handle of *PCP* or left (or bottom) handle in *SPLOM*, the selected axis/column are responsive and the counterpart elements on the other view turn to green. *Step 2:* User drag selected axis/column to desired location. In *PCP*, axis switch simultaneously during the dragging while in *SPLOM*, matrices are only reordered when user finish dragging. *Step 3:* User release the mouse and both views updated. *SPLOM* rows, columns, and cells updated with animated transition.

PCP Each data item is encoded as a connected polyline and the attributes as the intersection points of these lines with the axes (see Figure 1 (a) (left)). The polylines in *PCP* give primacy to individual data points and their data values across dimensions. Labels name the data dimensions of the axes and also the data ranges.

SPLOM We use the empty cells on the diagonal of the *SPLOM*(where, otherwise, each dimension would be plotted against itself) for the dimension labels. Data ranges are labelled at the left and bottom edges of the *SPLOM* (see Figure 1 (a)(right)).

Combined The *PCP* and *SPLOM* views were presented side-by-side, coordinated through brushing-and-linking as described below. (see Figure 1 (a)).

Our *Combined* view is intended to combine the advantages of both *SPLOM* and *PCP*. Interaction techniques can enhance the user's perception of information when visually exploring a dataset and reduce the drawbacks of the individual techniques—particularly those related to visual clutter and object overlap—providing the user with mechanisms for handling complexity in larger datasets [11]. The following list shows the interaction features we implemented on the combined views.

Item Identification In piloting different tasks for our study we found that interaction was essential to allow users to isolate and identify individual data items or to restrict the display to groups of data items. On mouse-hover, individual data lines (*PCP* view) and points (*SPLOM* view) are highlighted across all plots/axes, and their labels displayed near the mouse cursor.

Range filtering In order to limit line clutter *PCP* view must allow the user to filter the displayed data lines by data attribute ranges. This is achieved in *PCP* by allowing the user to select a section of an axis through mouse dragging [32]. Multiple selections on multiple axes are applied together, such that only data elements with lines passing through all the selected regions are highlighted. In the *SPLOM* view lasso select highlights data points within the rectangular selection across all plots [37]. Thus, *SPLOM* allows the user to filter in two data dimensions simultaneously.

Combined-view Brushing In coordinated-multiview displays it is fairly standard that items selected in one view are also highlighted across all other views – an interaction commonly known as *brushing*. In our combined *PCP-SPLOM* view, brushing is used such that when lines in the *PCP* view are selected or hovered-over, the points for the corresponding data items are also highlighted in the *SPLOM* view. Conversely, selections or hovers in the *SPLOM* view are also highlighted in the *PCP* view. Coordinated brushing of selections and individual points is shown in Figures 1 (a) and (b), respectively.

Axes and Matrix Reordering In *PCP* display, it is difficult to determine if a pair of data dimensions are strongly correlated unless the corresponding axes are adjacent. Thus, interactive reordering of axes by direct drag and drop is a fairly standard interaction for *PCP* systems [3]. Less standard is interactive reordering of *SPLOM* displays – indeed we are not aware of any visualisation systems that currently implement this feature over a *SPLOM*. However, to keep the two views coordinated during reordering of *PCP* axes, it was essential to allow reordering of the *SPLOM*, and furthermore, to ensure that interaction was reciprocal across both displays. That is, for symmetry, we felt it was important to allow the user to initiate a reordering of both views by dragging an individual cell in the *SPLOM*. Figure 2 illustrates this reordering interaction. To reinforce the coupling, both views are updated smoothly and simultaneously throughout the drag operation, regardless of in which view the reordering was initiated. Please see our video to see this effect in action.

4 STUDY

4.1 Hypotheses and Tasks

Based on our research questions stated in the introduction, we formulated the following hypotheses:

- H_0: There is no difference in time and error across techniques, for each task (null-hypothesis).
- $H_{overall}$: *Combined* will be at least as accurate as the best of the two individual techniques *PCP* and *SPLOM*.

We identified six generic tasks that require participants to access the data at different levels of detail: from a single observation point to detecting patterns in large subsets of the data. There are many tasks that could be tested. The six tasks we have chosen for this study are based on past work in parallel coordinate and scatterplot matrix evaluation [2,19,24,27,28]. These tasks were also identified within the ten low-level analysis tasks that are important within information visualization [1]. The tasks were couched to the study participants in terms of a dataset giving student names and their marks in five subjects. The tasks and the expected relative difficulty of answering using *PCP* and *SPLOM* are:

Best-Performer: *Which student has the best average mark?* This task translates into finding the best performer from 100+ students. The target student is the best overall score. In *PCP*, participants had to search for the top students in each subject and quickly read their values and make an estimation of average mark. In *SPLOM*, participants had to find the data point at the top right corner of certain dimensions while comparing values in dimensions that represent four subjects. In *Combined*, we hypothesised that some participants would set a filter on the *PCP* view and confirm the values across all dimensions in the *SPLOM* view.

Subset-Tracing: *What is the name of the student of age 22 and whose mark in Science is lower than or equal to 75 and who has the best Math score?* This task translates into finding the candidate who satisfies the three given conditions and requires participants to be able to follow some points over several axes in the plot and also to read the value for specific points. In *PCP*, participants had to set filters on both Science and Age dimensions, then follow the lines across all other dimensions. In *SPLOM*, participants had to set a filter on Science in one of the plots, and then locate the best math candidate in the Age versus Math plot. In *Combined*, we expected participants to obtain an accurate answer from *SPLOM* but to set filter on *PCP* because its selection brush is much easier and more accurate to operate (i.e. the vertical axes in *PCP* present a larger target area for interaction than the plot areas in *SPLOM*).

Object-Comparison: *Over all subjects (Sci, Tech, Eng, Math) which highlighted student had most similar average score to the student highlighted in blue?* This task offered three candidate students highlighted in yellow and one target student highlighted in blue. Participants had to estimate average score of every candidate and compare them with the target. We expected *PCP* and *SPLOM* to have similar performance, while in *Combined*, we expected participants would start with one view and check in the other.

Outlier-Detection: *Find the student who has a much higher mark in Engineering than all other subjects (Choose the most obvious one).* This task requires the participant to detect outliers or other anomalies of the data with respect to the given measure. In *PCP*, participants need to identify which candidate has a sharp rise to their Engineering mark. In *SPLOM*, the participant needed to pick out the most obvious singular point among the Engineering related dimensions. In *Combined*, we expected participants to prefer *SPLOM* for outlier detection in the *Combined* condition, as a previous study found *SPLOM* effective for this task [30].

Correlation-Estimation: *Which two subjects are most positively correlated?* This task asks participants to assess the linear correlation between all pairs of subjects. This is straightforward in *SPLOM* but more difficult in *PCP* as participants must recognise the cross-line patterns for correlation and reorder the axes to check all correlation for all pairs. Harrison *et al.* [14] found experimentally that *SPLOM* was significantly better for identifying positive correlation than *PCP* and *PCP* was again demonstrated to be poor for this task recently by Kwon and Lee [20]. The question our study aims to answer is whether participants will be able to efficiently select the better view for correlation in *Combined*.

Cluster-Identification: *In which pair of dimensions are there most clearly two distinct clusters?* In this task, the participant was required to find groups of similar points based on visual features such as proximity of lines or line density in *PCP* and clustering of points in *SPLOM*. These clusters might be visually separated or overlapping; it depends on how the grouping was performed. Participants were asked to find the most distinct dimension. Similar to the *Correlation-Estimation* task, we expected *PCP* to perform badly and that in *Combined* participants would rely on the *SPLOM* view.

The tasks *Best-Performer*, *Subset-Tracing*, *Object-Comparison*require users to work with single data items to identify, compare and calculate values. the tasks *Outlier-Detection*, *Correlation-Estimation*, and *Cluster-Identification*require participants focus on identify overall trends and general patterns of the dataset.

4.2 Experimental Datasets

We generated synthetic data in order to ensure the generalisability of our results. We used 6 dimensions, i.e. four subjects plus the age and the name of the student, for the study. The same data sets were used in each of the visualisation conditions. A variety of measures were used to hide this so as to avoid a learning effect. In each condition we used different names for the students and the subject names were swapped between the data dimensions and the order of dimensions shuffled in the visualisation. Modulo these changes the data for each task was identical in each condition.

Two difficulty levels, *easy* and *difficult*, were created by adjusting the density of the data and the distance between the correct answer and the alternatives. Each dataset contained 100-200 observation points (100-120 points for *easy* and 180-200 points *difficult*). For *Best-Performer*, *Subset-Tracing*, and *Object-Comparison*, there was only a 10-30% difference between the correct answer and the closest alternative, while for easy datasets the difference was 30-100%. For the *Outlier-Detection* task, in *difficult* datasets at least four obvious candidate points needed to be considered by participants, whereas only two candidates points required inspection in the *easy* dataset. For *Correlation-Estimation*, in *difficult* datasets the correlation coefficients ranged from 0.4 to 0.9 and there was 0.1 difference between the correct answer and its closest alternative. The *easy* datasets correlation coefficient ranged from 0 to 0.9 with at least 0.2 difference between the correct answer and its closest alternatives. For *Cluster-Identification*, 2-5 clusters were generated using the DBSCAN method [35]. The *difficult* datasets had a 50% shorter reachability distance (the ε parameter of DBSCAN) and the minimum number of points in *difficult* clusters were 30% to 60% less than those of *easy* datasets. Two examples of tasks with different difficulty levels are shown in Figure 1.

We created 15 datasets for each trial: 3 small data sets for training, 6 easy datasets and 6 difficult datasets. Data was generated using R and Mockaroo [25]. Subject scores were generated from 35 to 100, and age was generated from 17 to 23. Names were randomly generated from a name database.

4.3 Participants and Setting

In order to avoid participant fatigue we split the experiment into two separate parts. Each contained three tasks. Participants were recruited for each part using ads on university-wide bulletin and email lists. Four $50 gift cards were offered as an incentive for top 30% of fastest and most accurate responses.

Participants: We recruited 21 participants (8 female, 13 male) for part A and 30 participants (9 female, 21 male) for part B, all with normal or corrected-to-normal vision and without colour vision impairment. The two parts of the experiment were conducted two weeks apart. 9 participants attended both experiments. The participants' age ranged from 19 to 55 with an average age of 29 years old. Participants were asked about their prior experience using *PCP* and

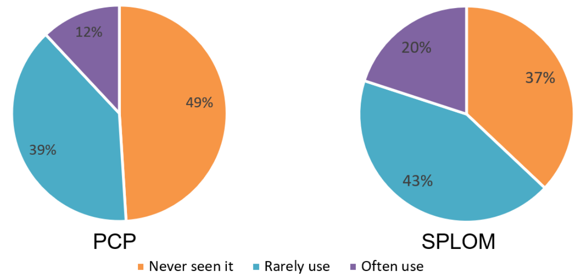

Figure 3: Self-declared prior experience of the 51 study participants.

SPLOM. Available options are: '*Never seen before*', '*Rarely use*' or '*Often use*'. As shown in Figure 3, 19 out of 51 participants (37%) had never seen *SPLOM* before, while 25 out of 51 participants (49%) had never seen *PCP* before. Participants were mostly students and staff from our university.

Environment: The study was run on an Intel Core i7 MacBook Pro (2016) and a 24-inch screen (1920 $\times$ 1080). During experiments, participant gaze was tracked with a Tobii X3-120 Eye tracker. The visualisation area was centred in a full-size window and participants interacted with it using a mouse and keyboard. Visualisation size was the same for all visualisations across all conditions.

4.4 Experimental Design & Procedure

Design: We used a within-subject, full-factorial design: 3 techniques $\times$ 3 tasks $\times$ 12 + 3 training trials for each part of the experiment. Since completing all six tasks required about 90 minutes, we split the study into two, approximately 45 minute parts, run with different participants (as detailed above) on different days to avoid fatigue. In Part A we used the first 3 tasks (*Best-Performer*, *Subset-Tracing*, and *Object-Comparison*) and the remaining 3 tasks (*Outlier-Detection*, *Correlation-Estimation*, and *Cluster-Identification*) in Part B. Part B was run about two weeks after Part A. We counter-balanced the techniques, creating two groups: one group started with individual view (*PCP* or *SPLOM*)(*trained*), while the other group started with *Combined* (*untrained*). In Part B, the 9 participants who participated in both studies were evenly distributed across these groups. We were interested if participants' use of the two views in *Combined* and their performance varied depending upon whether they were *trained* on both views individually before using *Combined*. Question order was randomised.

During each trial a timer progress bar was displayed on the top of the screen. Participants were instructed to remember the answer once they found it and to press the space bar to view the answer options. At that point, the timer stopped and the visualisation disappeared. Completion time was measured from the point when visualisation was shown on the screen, to the time when the participant pressed the space bar. Tasks timed out after 30 seconds, and the participants were shown the answer options. Answer options included (depending on the task) a set of candidate data item names or axes names as well as *"None of the above"* and *"Too difficult"*.

The experiments were conducted individually in a laboratory with the presence of one instructor.

After receiving information about the study through the explanatory statement and agreeing to the consent form the study took the following structure:

Background knowledge survey: Participants were asked to select their prior experience using *PCP* and *SPLOM* (see Figure 3).

Eye-tracker calibration: Participants were asked to run a calibration for the eye-tracker (took about 20 seconds).

Trials: The participants were then presented in turn with with the three visualisation conditions, and then for each of condition they were presented in turn with the three different tasks.

Training: At the beginning of each condition and task, the experimenter:

1. Instructed each participant on the task and how the visualisation might be used to complete the task.
2. Advised that they have three training examples for that task and visualisation technique, that the correct answers to the training examples are provided after they submit their answers, and that they can ask questions during the training.
3. Guided them through the first training example, demonstrating available interactions and encouraged the participant to practise these. Then told them to try the next two training examples. Participants were only allowed to proceed once they had correctly answered each of the training examples.
4. Advised them that they must now answer the trials on their own and that their answer and time would be recorded and that there was a 30 second time limit. They were asked to complete trials as accurately and quickly as possible.

Ranking and Feedback: At the end of the study participants were asked to rank the three visualisation methods in terms of their preferences to solve each of three tasks. They were also given opportunities to comment on the strengths and weaknesses of each visualisation method between tasks.

5 Results and Discussion

5.1 Data analysis

Data Collected: The following data was collected from each participant:

- prior knowledge about *PCP* and *SPLOM*,
- *completion time* of each trial,
- answer of each trial for *accuracy*,
- eye-tracking data and all keyboard/mouse events,
- user preference of visualisation for each task.

For each of the three tasks in part A, we obtained 12 trials $\times$ 3 techniques $\times$ 21 participants = 756 trials (2268 trials in total). For each of the three tasks in part B, we obtained 12 trials $\times$ 3 techniques $\times$ 30 participants = 960 trials per task (2880 trials in total).
Data treatment: We removed 6 trials that had a task-completion time less than 1 second as we considered them accidental clicks on the space-bar. We treated the answers '*Too difficult*' and '*None of above*' as errors, meaning that for each task the response was either *Correct* or *False*. Reported completion times are for correct answers only. We decided to keep trials where the participant had hit the time limit for each task as participants were still able to give an answer and we were more interested in *accuracy* than in *completion time*. We also checked whether this affected our results, rerunning the analyses after removing those timeout-trials, and found the same result with respect to significances and ranking of completion time and accuracy.
Statistical analysis: We analysed each of the tasks individually. Both *completion time* and *accuracy* measures were **not** normally distributed and we could not correct this by any standard transformation (Box-Cox transformation, log-transformation). We made sure that this was not an artifact of including trials that hit the time-limit. Time and accuracy were analysed individually for each task using the non-parametric Friedmans' test for one-way factorial analysis between techniques per test with a significance level of $p < 0.05$. For post-hoc pair-wise comparison, we used Mann Whitney U test, as removing trials resulted in unequal sample sizes.

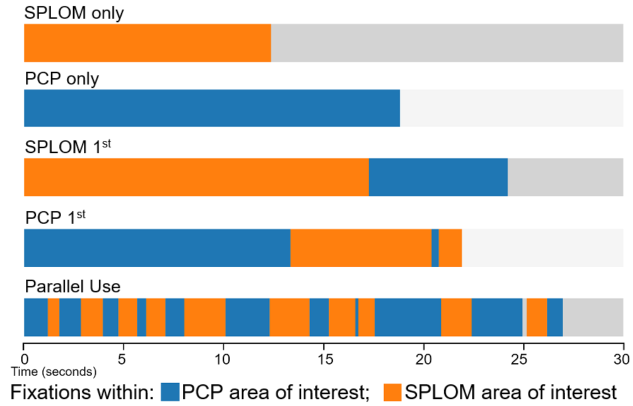

Figure 4: Five commonly used strategies on *Combined* views.

In the sections below, we report on results with a significance level of $p < 0.05$ (*), $p < 0.01$ (**), and $p \leq 0.001$ (***) , for each task individually. Numbers in brackets indicate mean values in seconds (time) and mean-errors.

Figures 5, 6 and 7 depict the main descriptive statistics for the three main objective measures, namely *accuracy* and *completion time*, and *user preference* rating.

5.2 Eye-tracking data

We were interested in analysing the eye-tracking data to understand the strategies employed when using the *Combined* view.
Eye-movement data extraction: The eye movement and mouse data was post-processed in Tobii Studio into separate fragments for each trial using the *Combined* view. We defined each of the two views in the *Combined* view as Areas of Interest (AOI) in order to study how participant attention was distributed between the views, assuming that they were attending to whichever view they were looking at. For each trial in *Combined* we measured the fixation and visit duration participants spent looking at each of the two AOIs. Then the fixation data based on the AOIs was computed and video clips containing overlaid gaze paths were produced and used for checking and identifying participant strategies, before applying statistical techniques across AOI fixation data to categorise participants by the identified strategies as detailed below. Figure 10 shows the aggregated relative visit duration that all participants spent on each view in *Combined* view.
Eye-movement data analysis: Since Tobii studio has limited visualisation capacity, we extracted fixation and visit duration data and created our own visualisations to show the eye movement for each trial and visualised them together with participants' accuracy results (see Figure 9). We then analysed the eye-tracking path videos and data individually for each participant and identified five common strategies (see Figure 4). To deal with the noisiness of eye-tracker data, null fixation duration less than 200ms between two same AOI segments were treated as the same AOI; The minimum fixation duration filter is set to 100ms; trials that including total null fixation time larger than 5 seconds were excluded. The following classifications were applied after eliminating any noise fixations, as follows:

SPLOM Only – participants stick to *SPLOM* only. Any fixations on *PCP* last no more than 1 second.

PCP Only – participants stick to *PCP* only. Any fixations on *SPLOM* last no more than 1 second.

SPLOM 1st – participants use *SPLOM* first, then switch to *PCP* (Sequential use), may occasionally have more switches but for no more than 1 second.

PCP 1st – participants use *PCP* first, then switch to *SPLOM* (Sequential use), may occasionally have more switches but for no more than 1 second.

Parallel Use – participants switch frequently between both views.

Figure 9 is an example eye-tracking visualisation from the study: the participant in (a) uses *Parallel Use* strategy that frequently switches between views to solve tasks (Tasks 1-3 were training, and tasks 4-15 are trials). Percentages of view use and results are displayed on the right for each trial. Please see our video of eye-movement recording of common strategies. Participant in (b) is an example use of the *PCP 1st* strategy.

5.3 Task results

***Best-Performer*:** For *completion time*, no statistical significance was found between the visualisation techniques. For *accuracy*, FRIEDMANS' TEST found *PCP* and *Combined* (0.88) are significantly (**) more accurate than *SPLOM* (0.81). This result supports $H_{overall}$ and contradicts H_0. We expected participants to spend more time on *PCP* since novice users were not familiar with parallel coordinates may take extra time to learn unfamiliar visualisation. Eye-tracking data confirmed this, with participants spending more time on *PCP* than *SPLOM* in combined view. *PCP 1st* (41%) were most popular and *Parallel Use*(29%) were second popular, and no one use *SPLOM Only* in *Combined* views. (See Figure 8 Left.)

***Subset-Tracing*:** For *completion time*, we found *PCP* (16.3sec, SD=6.2) and *Combined* (15.6sec, SD=5.8) to be significantly (**) faster than *SPLOM* (21.2sec, SD=6.1). We expected *SPLOM* to be slower and less accurate due to increased cognitive effort in tracing and filtering subset. For *accuracy*, we found *SPLOM* (0.63) was significantly (***) worse than the other two views (*Combined*: 0.94, *PCP*: 0.92). Again, Eye-tracking data shows that participants spend more time on *PCP* in the combined view. Figure 8 reveals that *PCP Only* was most popular (41%), with the rest using both views to check or verify their answers *SPLOM 1st* (18%), *PCP 1st* (18%), and *Parallel Use* (24%), Again, no one use *SPLOM Only* in *Combined* views..

Measure	Tech.	Best-Performer	Subset-Tracing	Object-Comparison	Outlier-Detection	Correlation-Estimation	Cluster-Identification
Time	**C**	16.8	16.3	14	14	8.4	9.2
	PCP	16.5	**15.6**	13.4	19.8	20.4	13.2
	SPLOM	19.4	21.2	16.3	**11.3**	**6.8**	7.8
Accuracy	**C**	**0.88**	**0.94**	**0.83**	**0.92**	0.84	**0.98**
	PCP	**0.88**	**0.92**	**0.81**	0.79	0.65	0.88
	SPLOM	0.81	0.64	0.62	**0.91**	**0.93**	**0.97**
User Preference	**C**	0.38	0.29	**0.62**	**0.57**	0.4	0.43
	PCP	**0.43**	**0.67**	0.24	0.2	0.07	0.03
	SPLOM	0.19	0.04	0.14	0.23	**0.53**	**0.53**

Figure 5: Mean results broken down by task for each of the three visualisation techniques. Results show time, accuracy and percentage of users preferring that visualisation. Yellow background indicates values that are significantly different from the other two. Statistically significant best values are highlighted in bold.

***Object-Comparison*:** For *completion time*, we found no statistically significant difference between the visualisations. For *accuracy*, we found *SPLOM* (0.61) to be significantly (***) less accurate than *PCP* (0.80) and *Combined* (0.81). Again, with no difference in completion time and better accuracy, this supports $H_{overall}$ and contradicts H_0. Eye-tracking data showed *Parallel Use* (35%) is the most popular strategy, and an equal porpotion of participants use both *PCP 1st* (29%) and *PCP Only* (29%), no one using *SPLOM Only* in *Combined* views.

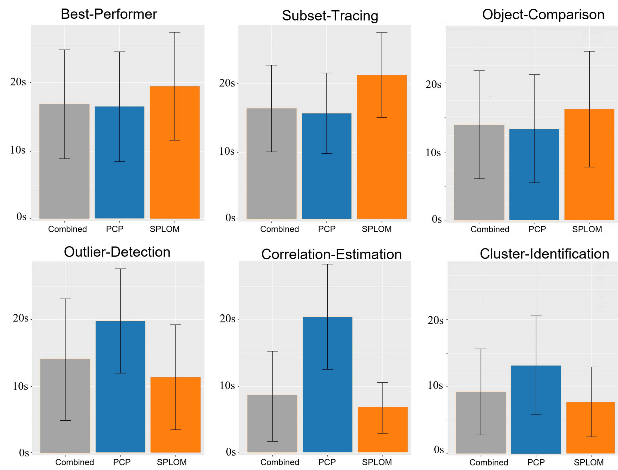

Figure 6: Mean and standard deviation for *completion time* broken down by task for the three visualisation techniques in *Combined* view across six tasks.

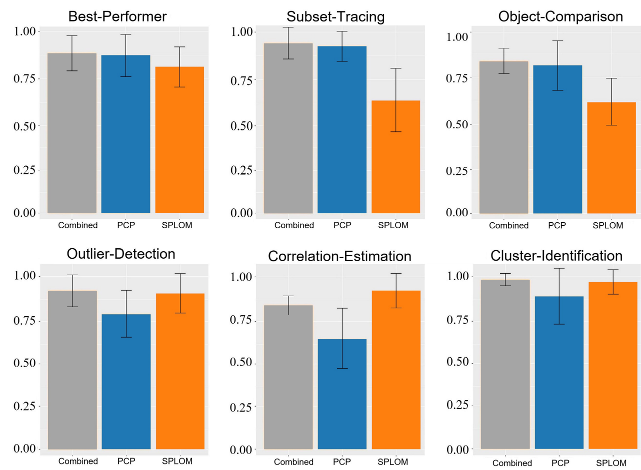

Figure 7: Mean and standard deviation for *accuracy* broken down by task for the three visualisation techniques in *Combined* view across six tasks.

***Outlier-Detection*:** For *completion time*, we found that *PCP* (19.7sec, SD=7.8) was significantly (***) slower than *SPLOM* (11.3sec, SD=7.8) and *Combined* (13.9, SD=9.1). For *accuracy*, we found that *SPLOM* (0.90) and *Combined* (0.92) to be significantly (***) more accurate than *PCP* (0.13). This supports $H_{overall}$. Eye-tracking data shows that *Parallel Use* (27%) and *SPLOM 1st* (27%) are both the most popular choice, and *PCP Only* (12%) and *PCP 1st* (12%) were lease popular. This result suggests that *SPLOM* was the most favoured choice for *Outlier-Detection*.

***Correlation-Estimation*:** For *completion time*, we found *SPLOM* (6.7sec, SD=3.8) and *Combined* (8.4sec, SD=6.7) to be significantly (***) faster than *PCP* (20.4sec, SD=7.8). For *accuracy*, we found significance (***) between all techniques: *SPLOM* (0.93) was the most accurate, followed by *Combined* (0.84), while *PCP* (0.65) was the least accurate. We expected *SPLOM* to perform better than *PCP* since it provides a straightforward view of all pairs of dimensions whereas with *PCP* the participant has to change the order of axes to check all possible combinations. Eye-tracking data shows that no one used a *PCP Only* strategy;

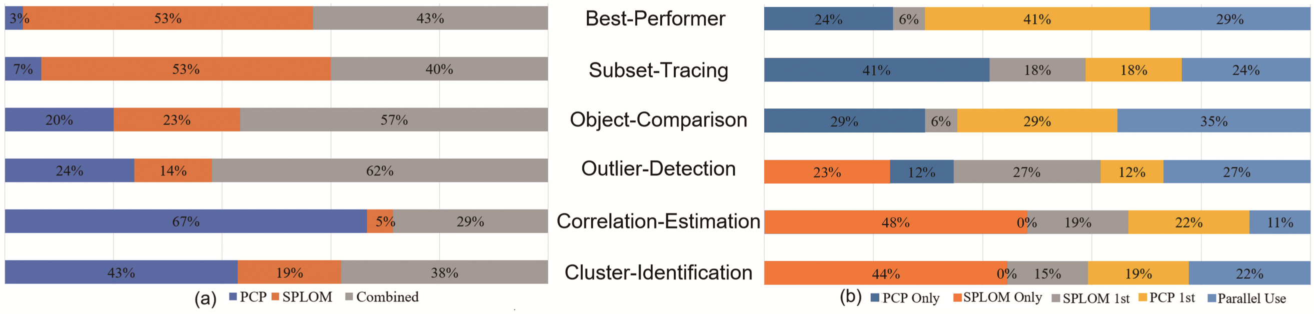

Figure 8: *Left:* (a):User preference of view in *Combined* view for each task. *Right:* (b) Strategies used in *Combined* view at each task (Based on aggregated eye-tracking visit duration data)

SPLOM Only (48%) is most popular; and the rest used both: *SPLOM 1st* (19%), *PCP 1st* (22%), *Parallel Use* (11%). No participant used *PCP Only*. This was the only task in which *Combined* was found to be less accurate than the generally superior of the two independent views (*SPLOM*). This result does not support $H_{overall}$.

***Cluster-Identification*:** For *completion time*, we found *SPLOM* (7.7sec, SD=5.2) is significantly (***) faster than *PCP* (13.2sec, SD=7.4) and *Combined* (9.2sec, SD=6.4). For *accuracy*, we found *Combined* (0.98) was most accurate, followed by *SPLOM* (0.97), and *PCP* was the least accurate. This result supports $H_{overall}$. During the training session, participants spent a larger amount of time on *PCP*, where eye-tracking data shows that *SPLOM Only* (44%) is the most popular strategy, while the rest used *Combined* for cross-checking: *SPLOM 1st* (19%), *PCP 1st* (22%), and *Parallel Use* (22%). Again, no participant used *PCP Only*.

View Use: Based on the eye-tracking data visualised as shown in Figure 4, we determined the time spent looking at each view in each trial using the *Combined* visualisation and categorised the strategy used. The aggregated results are shown in Figure 10 and 8. We also compared strategies of participants who were shown the individual views first (*trained in single-views*) to those shown *Combined* views first (*untrained in single-views*). We found:

Over the six tasks, 63% participants on average chose to use both views and 37% chose to use one view only.

Most of the participants *trained in single-views* used *PCP 1st* and *Parallel Use* in *Best-Performer*, *Subset-Tracing*, and *Object-Comparison*, and *SPLOM 1st* and *Parallel Use* in *Outlier-Detection*, *Correlation-Estimation*, and *Cluster-Identification*.

Participants *untrained in single-views* were more likely to use both views. They were very likely to choose *Parallel Use* as their strategy, while participants *trained in single-views* tended to choose a mix of strategies. We suspect that this was because the untrained participants were not sure which view was most appropriate for the task and were still exploring which view is better to help them solve the task.

5.4 Overall User Preference

After each task, participants were asked to choose a single visualisation technique that they preferred to use to solve the task. We used a CHI-SQUARED TEST for this data analysis. There was no statistically significant preference for the *Best-Performer* task. *PCP* was most popular for *Subset-Tracing*, and *Combined* most favoured in the *Object-Comparison* (62% (***)) and *Outlier-Detection* (57% (**)) task, while *SPLOM* (67% (***)) was most favoured for *Correlation-Estimation* (53% (***)) and *Cluster-Identification* (53% (***)). On the other hand, *SPLOM* was the least preferred for the *Subset-Tracing* and *Object-Comparison* tasks, while *PCP* was the least preferred choice for *Correlation-Estimation* and *Cluster-Identification*.

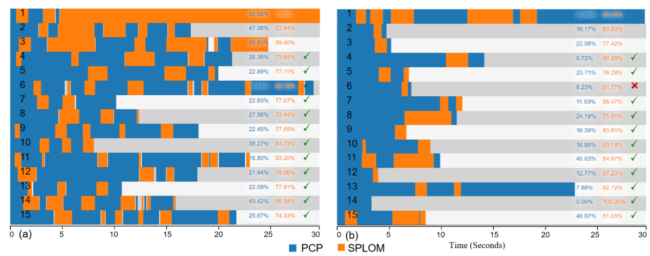

Figure 9: *Upper (a)*: An example of *Parallel Use* strategy for 15 trials, training tasks (1-3) only display first 30 seconds. *Lower (b)*: An example use of *PCP 1st* strategy. Strategy classification rationale is described in Section 5.2. Percentages in blue and orange represents time used in *PCP* and *SPLOM*. Tick and cross represent correctness of the result (training tasks excluded).

5.5 Discussion

We can confirm that *Combined* was always as good as, or sometimes even slightly better than the best of the individual techniques (accepting $H_{overall}$). However, we do not find any difference in *completion time* for people using both views together or exclusively using one view in the *Combined* condition.

Participants were able to identify the best view for each task and swiftly change their strategies accordingly. Figure 8 shows that selected strategies varied between tasks.

Correlation-Estimation is the only task where *SPLOM* was better than *Combined* in terms of both *accuracy* and *completion time*. The result that people struggle to identify correlation using *PCP* confirms earlier studies [14,20]. In *PCP* participants have to re-order the axes in order to examine all dimensions whereas in *SPLOM* it is straightforward to identify correlation. In this case, the *PCP* view may became a distraction or confuse people, hence the accuracy was lower in the *Combined* view than in *SPLOM* only view – although, *Combined* was still significantly better in both speed and accuracy than *PCP* only. Thus, although we did observe that participants unanimously chose to complete this task using *SPLOM* (Figure 8(b)) and benefitted from it, the presence of the *PCP* view still had a detrimental (distractor) effect on completion time for this task.

Despite a minor trade-off in terms of completion time, we have good results for accuracy in *Combined*, especially for the most difficult and complex tasks.

Averaged over all six tasks, 45% of participants chose *Combined* as their preferred view. However, eye-tracking results show that 63% chose a strategy that involved both views. These results are in accordance with the measures we found for *completion time*.

We think that the *Combined* view has an increased visual complex-

Figure 10: Relative duration (bar-height) spent looking at each view in *Combined* view. Task 1-3 are Training tasks.

ity which may lead to longer completion time. However, except in *Correlation-Estimation*, we do not see any statistical significance in the increase of *completion time* between *Combined* and the superior view for the task.

We also found that novice users (according to self reported experience with the given visualisations) were only able to use basic complementariness such as cross-checking their answers. By contrast, some expert users seem able to interact with both views to effectively bypass the disadvantages of the worse view for a given task and hence benefit from the *Combined* view (e.g., in the *Object-Comparison* task, a notable number of participants performed selection/filtering on *PCP*, then used *SPLOM* to find answers, then returned to *PCP* for cross-checking).

In *Parallel Use* strategy, there were two groups of people: savvy users who actively utilised both views and novice users still gaining confidence with the individual visualisations. In the latter case, *Combined* view could be either a distraction or the two views together could have a positive, reinforcing and training effect. Resolving whether combined views support training is an interesting area for further study as discussed in Section 7.

In general, *SPLOM* are easier to use for the task of *Outlier-Detection*, *Correlation-Estimation*, *Cluster-Identification*, while *PCP* is the better choice for other tasks, including *Best-Performer*, *Subset-Tracing*, *Object-Comparison*.

We grouped participants by their strategies for each task and repeated our tests on those groups. The result accords with the earlier findings. In particular, *Parallel Use* and *PCP 1^{st}* or *SPLOM 1^{st}* have similar or even better results for *accuracy*, and *PCP Only* and *SPLOM Only* were identical with the single view performance.

5.6 Limitations

There are several limitations of our methodology that should be noted. First, although in 5 out of 6 tasks *Combined* views outperform single views, we found that the standard deviation is high and there are likely other advantages and disadvantages that affect individual users performance. In particular, there seems to be a great deal of individual difference in how effectively users are able to use the interactive aspects of our interface. For example, several users complain that highlighting & brushing, rather than supporting them, is a distraction. The data density used in the experiments are between 100-200 data records in each trail. As with all usability studies, it cannot be assumed that our results would apply to larger or differently structured data.

6 Conclusion

In this paper, we presented a combined representation of *PCP* and *SPLOM*, and explored and tested the effect of perceptually complementary views for multivariate data. We studied six tasks and performed a detailed analysis of participants' strategies for using the combined view based on eye-tracking data.

Within the limitations of our tested conditions, we can conclude the following:

***PCP* and *SPLOM* views are perceptually complementary.**
From the experiment above we can see that our complementary view has some clear advantages. In most multivariate data analysis tasks, showing data in a point distribution can greatly help users' comprehension overall (e.g., *Correlation-Estimation*, *Outlier-Detection*, and *Cluster-Identification*) and line visualisation can help them identify single observation points easily across all dimensions. With the *Combined* view, both data trends and detail tasks can be spotted easily.

Placing perceptually complementary views side-by-side is an effective way to get the benefits of both.
As described in Section 2, there have been many proposals for more complex interactive combinations of *PCP* and *SPLOM* views into a single interactive display. These combinations pose a significant design and implementation challenge and so far do not seem to have been tested with real users. That is, evaluation is based on qualitative use-case analyses by the authors and it is unclear whether other users would experience the same benefits. In the mean-time, what seems to us the simplest strategy for combining these views (placing them side-by-side) has been, to the best of our knowledge, hitherto untested. We do not claim that there is no benefit to these more complex combinations - but hopefully our results can provide a baseline upon which to improve.

Five common view use strategies were identified for *Combined* views based on eye-tracking data.
While it seems a majority of users obtained benefit from the combined view, it is important to note that not all users experienced the benefits. Our finding that more expert users used both views in the combined view leads us to speculate that with more training other users may also obtain benefits from the *Combined* view.

Some people use complementariness without awareness.
User preference data does not completely agree with eye-tracking results. Despite only 45% of participants voting for the *Combined* view as their favourite strategy, eye-tracking shows that 63% of tasks were completed by use of both views.

When one view is clearly better than the other for a particular task, there is a small but significant overhead in terms of completion time in *Combined* views.
Despite the complementary effects described above, when a particular view is overwhelmingly better for a given task (e.g. *Correlation-Estimation*), we did observe a relatively small but in some cases significant increase in completion time to the combined view over the best view for that task alone. Thus—while the combined view may be the best choice in complicated analysis involving multiple different subtasks, or frequent switching between tasks—it may be contra-indicated when only a single task is being performed repeatedly and there is a clearly better view for that task.

7 Future Work

We plan to continue studying the effects of complementariness on other types of datasets (e.g., for hierarchical data and dynamic networks). This can be done in a unified manner to maintain consistency across all data types. We would like to better understand which perceptually complementary views can improve *completion time* and *accuracy* and how to measure the degree of perceptual complementariness of two representations and possible learning effect when using combined views. In the paired views that have been studied so far (*PCP* and *SPLOM* in this work, different graph visualisations in [5]), there is a certain duality in terms of one view being data item centric (e.g. *PCP* and nodelink diagrams) and one being relationship centric (e.g. *SPLOM* and adjacency matrices). Is this a useful categorisation that suggests further candidate views for pairing in a perceptually complementary manner?

Most participants in our study seemed to able to use the combined view to good effect, particularly those who had prior experience with these kinds of visualisations. However, there was a sizeable minority who did not experience this benefit. We would like to study further the effect of training on people's effectiveness in using perceptually complementary displays, for example, to determine if the use of a combined display helps people to learn to use the individual views more effectively. However, this would require a different, longitudinal study design.

8 Acknowledgements

This research was supported under Australian Research Council's Discovery Projects funding scheme (project number DP140100077).

References

[1] R. Amar, J. Eagan, and J. Stasko. Low-level components of analytic activity in information visualization. In *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on*, pp. 111–117. IEEE, 2005.

[2] G. Andrienko and N. Andrienko. Constructing parallel coordinates plot for problem solving. In *1st International Symposium on Smart Graphics*, pp. 9–14, 2001.

[3] M. Ankerst, S. Berchtold, and D. A. Keim. Similarity clustering of dimensions for an enhanced visualization of multidimensional data. In *Information Visualization, 1998. Proceedings. IEEE Symposium on*, pp. 52–60. IEEE, 1998.

[4] I. Borg and P. J. Groenen. *Modern multidimensional scaling: Theory and applications*. Springer Science & Business Media, 2005.

[5] C. Chang, B. Bach, T. Dwyer, and K. Marriott. Evaluating perceptually complementary views for network exploration tasks. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1397–1407. ACM, 2017.

[6] J. H. Claessen and J. J. Van Wijk. Flexible linked axes for multivariate data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2310–2316, 2011.

[7] W. C. Cleveland and M. E. McGill. *Dynamic graphics for statistics*. CRC Press, Inc., 1988.

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

[9] M. Cordeil, A. Cunningham, T. Dwyer, B. H. Thomas, and K. Marriott. Imaxes: Immersive axes as embodied affordances for interactive multivariate data visualisation. In *ACM UIST*, 2017.

[10] A. Dasgupta and R. Kosara. Pargnostics: Screen-space metrics for parallel coordinates. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1017–1026, 2010.

[11] M. F. De Oliveira and H. Levkowitz. From visual data exploration to visual data mining: a survey. *IEEE Transactions on Visualization and Computer Graphics*, 9(3):378–394, 2003.

[12] U. Engelke, J. Vuong, and J. Heinrich. Visual performance in multidimensional data characterisation with scatterplots and parallel coordinates. *Electronic Imaging*, 2016(16):1–6, 2016.

[13] M. Ghoniem, J.-D. Fekete, and P. Castagliola. A comparison of the readability of graphs using node-link and matrix-based representations. In *Information Visualization, 2004 (InfoVis)*, pp. 17–24. IEEE, 2004.

[14] L. Harrison, F. Yang, S. Franconeri, and R. Chang. Ranking visualizations of correlation using weber's law. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1943–1952, 2014.

[15] J. Heinrich, J. Stasko, and D. Weiskopf. The parallel coordinates matrix. *EuroVis–Short Papers*, pp. 37–41, 2012.

[16] P. J. Huber. Projection pursuit. *The annals of Statistics*, pp. 435–475, 1985.

[17] A. Inselberg. The plane with parallel coordinates. *The visual computer*, 1(2):69–91, 1985.

[18] J. Johansson and C. Forsell. Evaluation of parallel coordinates: Overview, categorization and guidelines for future research. *IEEE Transactions on Visualization and Computer Graphics*, 2016.

[19] R. Kanjanabose, A. Abdul-Rahman, and M. Chen. A multi-task comparative study on scatter plots and parallel coordinates plots. In *Computer Graphics Forum*, vol. 34, pp. 261–270. Wiley Online Library, 2015.

[20] B. C. Kwon and B. Lee. A comparative evaluation on online learning approaches using parallel coordinate visualization. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 993–997. ACM, 2016.

[21] M. Lanzenberger, S. Miksch, and M. Pohl. Exploring highly structured data: a comparative study of stardinates and parallel coordinates. In *Information Visualisation, 2005*, pp. 312–320. IEEE, 2005.

[22] J. Li, J.-B. Martens, and J. J. Van Wijk. Judging correlation from scatterplots and parallel coordinate plots. *Information Visualization*, 9(1):13–30, 2010.

[23] S. Liu, D. Maljovec, B. Wang, P.-T. Bremer, and V. Pascucci. Visualizing high-dimensional data: Advances in the past decade. *IEEE Transactions on Visualization and Computer Graphics*, 23(3):1249–1268, 2017.

[24] L. Micallef, G. Palmas, A. Oulasvirta, and T. Weinkauf. Towards perceptual optimization of the visual design of scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 23(6):1588–1599, 2017.

[25] L. Mockaroo. Stichwort: data generator. *Online: http://www. mockaroo. com*, 2016.

[26] L. Nowell, R. Schulman, and D. Hix. Graphical encoding for information visualization: an empirical study. In *Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on*, pp. 43–50. IEEE, 2002.

[27] G. Palmas, M. Bachynskyi, A. Oulasvirta, H. P. Seidel, and T. Weinkauf. An edge-bundling layout for interactive parallel coordinates. In *Visualization Symposium (PacificVis), 2014*, pp. 57–64. IEEE, 2014.

[28] W. Peng, M. O. Ward, and E. A. Rundensteiner. Clutter reduction in multi-dimensional data visualization using dimension reordering. In *Information Visualization, 2004*, pp. 89–96. IEEE, 2004.

[29] J. C. Roberts. State of the art: Coordinated & multiple views in exploratory visualization. In *CMV'07. Fifth International Conference on*, pp. 61–71. IEEE, 2007.

[30] D. M. Rocke and D. L. Woodruff. Identification of outliers in multivariate data. *Journal of the American Statistical Association*, 91(435):1047–1061, 1996.

[31] H. Siirtola. Combining parallel coordinates with the reorderable matrix. In *Coordinated and Multiple Views in Exploratory Visualization*, pp. 63–74. IEEE, 2003.

[32] H. Siirtola and K.-J. Räihä. Interacting with parallel coordinates. *Interacting with Computers*, 18(6):1278–1309, 2006.

[33] J. W. Tukey and P. Tukey. Computer graphics and exploratory data analysis: An introduction. In *Proceedings of the Sixth Annual Conference and Exposition: Computer Graphics 85*, 1985.

[34] C. Viau, M. J. McGuffin, Y. Chiricota, and I. Jurisica. The flowvizmenu and parallel scatterplot matrix: Hybrid multidimensional visualizations for network exploration. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1100–1108, 2010.

[35] W.-T. Wang, Y.-L. Wu, C.-Y. Tang, and M.-K. Hor. Adaptive density-based spatial clustering of applications with noise (dbscan) according to data. In *Machine Learning and Cybernetics (ICMLC), 2015 International Conference on*, vol. 1, pp. 445–451. IEEE, 2015.

[36] L. Wilkinson, A. Anand, and R. Grossman. High-dimensional visual analytics: Interactive exploration guided by pairwise views of point distributions. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1363–1372, 2006.

[37] J. Yang, W. Peng, M. O. Ward, and E. A. Rundensteiner. Interactive hierarchical dimension ordering, spacing and filtering for exploration of high dimensional datasets. In *Information Visualization, 2003. INFOVIS 2003. IEEE Symposium on*, pp. 105–112. IEEE, 2003.

[38] X. Yuan, P. Guo, H. Xiao, H. Zhou, and H. Qu. Scattering points in parallel coordinates. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1001–1008, 2009.

[39] J. Zhang and D. A. Norman. Representations in distributed cognitive tasks. *Cognitive science*, 18(1):87–122, 1994.