P-Lite: A Study of Parallel Coordinate Plot Literacy

Elif E. Firat[a], Alena Denisova[b], Max L. Wilson[a], Robert S. Laramee[a]

[a]*University of Nottingham, UK,*
[b]*University of York, UK,*

Abstract

Visualization literacy, the ability to interpret and comprehend visual designs, is recognized as an essential skill by the visualization community. We identify and investigate barriers to comprehending parallel coordinates plots (PCPs), one of the advanced graphical representations for the display of multivariate and high-dimensional data. We develop a parallel coordinates literacy test with diverse images generated using popular PCP software tools. The test improves PCP literacy and evaluates the user's literacy skills. We introduce an interactive educational tool that assists the teaching and learning of parallel coordinates by offering a more active learning experience. Using this pedagogical tool, we aim to advance novice users' parallel coordinates literacy skills. Based on the hypothesis that an interactive tool that links traditional Cartesian Coordinates with PCPs interactively will enhance PCP literacy further than static slides, we compare the learning experience using traditional slides with our novel software tool and investigate the efficiency of the educational software with an online, crowdsourced user-study. User-study results show that our pedagogical tool positively impacts a user's PCP comprehension.

Keywords: Human-Centered Computing, Visualization Systems and Tool, Empirical Studies in Visualization

1. Introduction

Interpretation and comprehension of data visualizations is gaining importance due to an ever-increasing volume of data. The use of visual designs

---

for exploratory analysis of large and complex data requires visualization literacy skills for non-expert users. Users' ability to interpret visualizations may vary depending on how visual representations present the data (plain or embellished) [1] and their own background and experiences [2]. Visualization literacy is studied extensively in the visualization community [3, 4, 5]. In the keynote talk at the IEEE VIS'19 conference, Börner [6] identifies visualization literacy as an essential skill, as crucial as general literacy used in everyday life. Two workshops, one at EuroVis'14 *"Towards Visualization Literacy"* [7], and another at IEEE VIS'14 *"Towards an Open Visualization Literacy Testing Platform"* [8] indicate that visualization literacy is gaining in popularity in the visualization community. Visualization literacy is defined as, *"The ability to confidently use a given data visualization to translate questions specified in the data domain into visual queries in the visual domain, as well as interpreting visual patterns in the visual domain as properties in the data domain"* [3]; *"The ability to make meaning from and interpret patterns, trends, and correlations in visual representations of data"* [9]; *"The ability and skill to read and interpret visually represented data in and to extract information from data visualizations"* [10].

Parallel coordinates plots (PCPs) provide a graphical representation of multidimensional relationships through the use of parallel axes (see Figure 1 and 2). This design can display high-dimensional data with up to 10-15 dimensions in practice, as each axis is visually separated [11]. Each polyline represents a data record that intersects the parallel axes at given points that indicate the value of individual dimensions. In comparison to a Cartesian Coordinate Plot (CCP), for example, PCPs display this multidimensional data in a plane that offers additional advantages (see Figure 2). The process of plotting data is different in the CCP and PCP spaces.

In this paper, we define **PCP literacy** as the ability to correctly read, interpret, and construct PCPs. PCP literacy is essential for any user who is interested in understanding multidimensional data, as this is what separates PCPs from other more common visual designs. PCPs, however, have a reputation of being difficult to comprehend, called an expert-only visual design, especially if the implementation lacks essential features e.g. interaction [13].

We introduce a novel literacy test to evaluate and improve PCP literacy. In the test, we include datasets and images generated using popular, off-the-shelf PCP tools. Based on our experience of teaching PCPs in the classroom, we develop an interactive pedagogical tool that advances PCP literacy skills by enabling novices to enhance their comprehension, interpre-
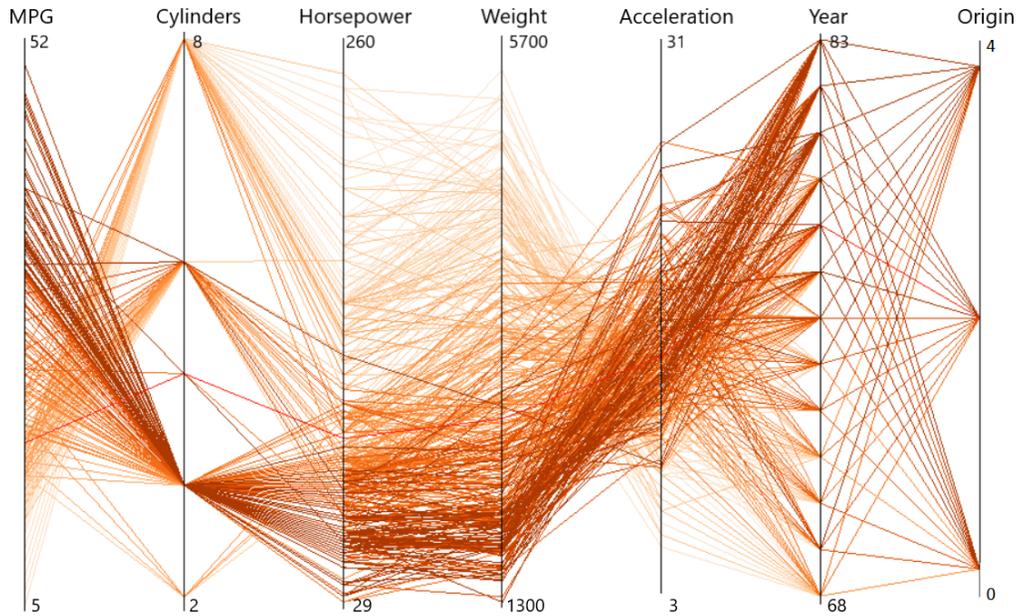
2

Figure 1: An example parallel coordinate plots of car dataset with 7 attributes. The image was created using Xmdv [12].

tation, and construction of PCPs. Besides supporting literacy skills of users, the tool empowers the effective transformation of data into knowledge and can be used to support an active learning experience in the classroom. Our fundamental hypothesis is that a software tool that interactively links CCP and PCPs will advance PCP literacy more than static slides alone. We assess the learning experience using traditional slides versus our novel software tool and investigate the efficiency of the educational software on PCP literacy with an online, crowdsourced user-study. The main contributions of this paper are:

1. A discussion of the barriers to PCP literacy (see Section 3);
2. Developing a novel educational tool that facilitates both the teaching and learning of parallel coordinates, advancing parallel coordinates literacy [14] (see Section 5);
3. Developing a novel PCP literacy test [15] and conducting a user-study to assess the impact of our pedagogical software on the comprehension of parallel coordinates (see Section 4 and 6).
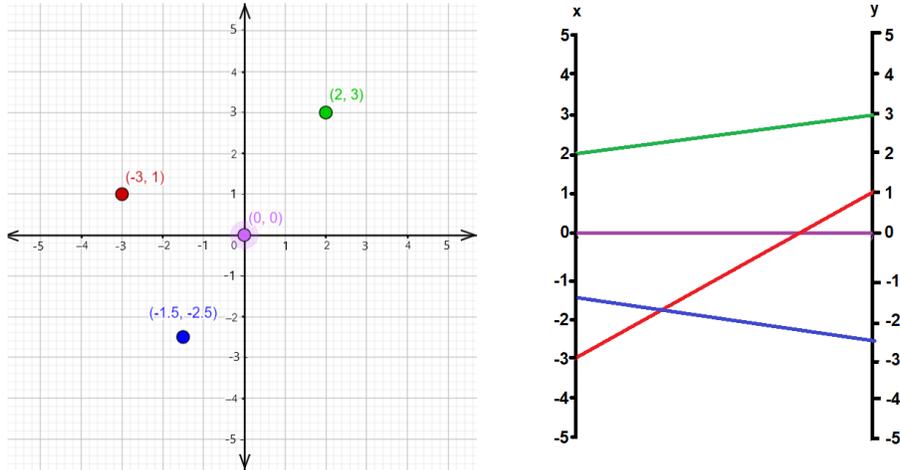
Figure 2: An example of Cartesian Coordinate Plot and Parallel Coordinate Plots with a 2D point data.

Supplementary Material	URL
Educational PCP Software [14]	https://bit.ly/3ddTMJl
Instructions for the Experiment [16]	https://bit.ly/36UobZH
Slides Video Tutorial [17]	https://bit.ly/36MSRvU
Software Video Tutorial [18]	https://bit.ly/3ix3ZSZ
PCP Literacy Test [15]	https://bit.ly/3xOUEMr

Table 1: The table summarizes the supplementary materials for our PCP literacy test with URLs.

Table 1 provides a summary of supplementary material for this PCP literacy study. The supplementary material makes the study fully reproducible.

The rest of the paper is organized as follows: Section 2 introduces related work that includes related literature on visualization literacy and PCPs with a user-study evaluation. Section 3 identifies and presents some of the challenges of interpreting parallel coordinates. The development of a parallel coordinates literacy test is presented in Section 4 and the educational literacy tool is introduced in Section 5. Section 6 explains how the user-study was conducted and Section 7 analyses the results. Discussion is introduced in Section 8. Finally, a conclusion is presented in Section 9.

2. Related Work

There are several previous projects on visualization literacy that study the influence of visual designs on a user's understanding and advance comprehension of visual interfaces. Some studies investigate the user's visualization literacy skills by providing reviews of user-study results. We searched the Survey of Surveys (SoS) on information visualization [19], surveys on visualization literacy [20] and interactive visualization for education [21], and a study of information visualization books [22] for related literature on visualization literacy. In this section, we present a collection of visualization literacy research and parallel coordinate papers that include user-studies with PCPs. While visualization literacy papers focus on investigation and improving to novice user's literacy skills, papers with user-studies evaluate PCPs to inform and evaluate the usability of the design. Differences between previous studies and the work presented here are summarized in Table 2.

2.1. Visualization Literacy

Börner *et al.* [9] investigate the general familiarity of various visual designs to museum visitors. Some 20 textbook visualizations and commonly used online visualization libraries such as the D3.js [23] library were selected to perform research in three U.S. science museums, which is considered an informal learning setting. There are two charts, five maps, eight graphs, and five network layouts in these visual models. Five of the 20 graphic representations were shown to science museum visitors. Visitors to the museum were asked to state their familiarity with the models and to identify the design name.

Focusing more on children, Alper *et al.* [4] explore visualization literacy techniques for elementary school children and develop an internet platform called C'est la Vis, which enables users to generate and interact with visual designs. It is used by classroom teachers to teach visualization by developing exercises for children such as reading, completing, and creating visualizations. We also look at communicating visualizations to learners, but focus more on a specific type of visualization (PCPs) rather than on teaching to a young demographic.

The study by Ruchikachorn and Mueller [24] introduces a learning-by-analogy technique that shows a step-by-step transformation between two visual models to explain an unfamiliar visualization method. The concept is illustrated using four examples of visualization pairs such as a data table

and parallel coordinates, a scatterplot matrix and hyperbox, a linear chart and spiral chart, and a hierarchical pie chart and treemap. After interacting with the transitions, the users can understand the new visual models more quickly. Similarly, the main focus of our work is to display the correspondence between a CCP and PCP with interactive animation to explain the PCP design concept better. We also develop a special-purpose PCP literacy test in conjunction with the pedagogical tool developed.

The purpose of a study provided by Boy *et al.* [3] is to create a visualization literacy assessment technique. They use Item Response Theory (IRT) [25] to assess the relevance of visualization literacy test items and the abilities of the examinees. The primary aim is to produce quick, effective, and reliable tests that researchers can use to identify test takers with lower visualization literacy skills. The study explains the design and assessment of two line graph visualization literacy tests and provides an extension of the method to bar charts and scatterplots. In contrast, our study presents a test with a focus on developing PCP literacy. We investigate the PCP literacy skills of participants and what they struggle with the most when answering questions related to multi-dimensional data.

Lee *et al.* [10] develop a test to evaluate the visualization literacy skills of users who are non-experts in data visualization. The test, called VLAT, consists of 12 visual designs and 53 test items featuring eight main tasks. A pilot study validates the test material and its reliability. They also study the connection between the VLAT results of users and their ability to understand an unfamiliar visualization. Conceptually, our work extends that of Lee *et al.* by adding a further more advanced visual design, namely PCPs.

Research by Firat *et al.* [5] identifies barriers to treemap literacy with a view to enhancing the skills of non-expert users in treemap literacy. The outcomes from two years of information visualization assignments are used to identify obstacles and challenges to the comprehension and construction of treemaps. A treemap literacy test is developed to assess the user's understanding and ability to interpret treemaps. The study presents a pedagogical treemap tool to promote both the teaching and understanding of treemaps and the advancement of treemap design. Further, a classroom-based study with 25 participants was conducted to evaluate the effectiveness of the educational tool. Investigation of barriers to visualization literacy, the assessment test, and the educational tool presented in this study share similarities with our study except the focus here is on PCP literacy. In general, the work we present here differs from previous work on visualization literacy by including

multi- and high-dimensional visual designs in the evaluation. Previous work focuses on lower-dimensional imagery as reflected in Table 2.

2.2. User-Studies with Parallel Coordinates

Inselberg has written extensively on parallel coordinates and provides an in-depth guide [26]. We found three survey papers using the SoS [19] that concentrate on parallel coordinates. A survey by Dasgupta *et al.* [27] seeks to identify various sources of uncertainty in screen space and link them to different uncertainty effects on the user. They review the research on parallel coordinates and use a taxonomy to classify different techniques to reduce ambiguity [27]. Another survey provided by Heinrich and Weiskopf surveys the parallel coordinates literature and develops a categorization, aiming to guide research into new topic-related directions [28]. Johansson and Forsell present a detailed literature review that focuses on user-centered evaluation and analyses the usability of parallel coordinates. The goal is understanding how people use PCPs to identify barriers to PCP literacy as well as providing a set of guidelines for future studies [29].

Yang *et al.* [30] propose a system for interactive hierarchical displays (IHDs) to tackle the clutter challenge associated with hierarchical multivariate visualization (HPC) techniques in the study of large datasets. For the evaluation of IHDs, 20 users are asked to identify patterns using two separate forms of parallel coordinate plots: flat parallel coordinates (FPC) and HPC. Patterns have been broadly described as either clusters or outliers. Participants found 8-9 out of the 25 patterns that were used. Subjects using HPC are more successful than subjects using FPC in finding trends in large datasets. This work assesses whether the proposed framework of IHDs provides users with effective help in exploring large datasets. Their study assumes that participants already have a basic knowledge of PCPs whereas ours does not.

Siirtola [31] compares two functionally different visual designs: PCPs and the Reorderable Matrix (RM), and examines how the two visual designs can be combined. An experiment is conducted where 20 participants performed tasks with an application featuring RM and PCP views of the same data, with and without linking. Results indicate that while the view linking originally slows the performance of user tasks, it accelerates learning and is well received by users. Similarly, our study indicates links between CCPs and PCPs rather than RM and PCPs.

Research Paper	User-study Theme	Visual Design	Image Generation	Evaluation Method	Evaluation Tool	Participant	Task \ Question
Visualization Literacy							
Boy et al. [3]	Calibration of visualization literacy test	BC, LC, SP	Manual	Crowdsourcing	Amazon MTurk [32]	43	48T
Ruchikachorn and Mueller [24]	Testing new framework that links unfamiliar visualizations to familiar ones	LC, PC, S, SP, PCP, T	Processing[33]	Crowdsourcing	Amazon MTurk [32]	22	12Q
Börner et al. [9]	Determining the familiarity of users with different visual designs	BC, PC	Manual	In-field	Forms	273	100Q
Alper et al. [4]	Testing users' interests and understanding of their activity	BC, P	C'est La Vis (C)	Classroom	Tablets	21	7T
Lee et al. [10]	Evaluation of content validity and test reliability	BC, CM, H, LN, PC, SC, SP, T	Manual	Crowdsourcing	Amazon MTurk [32]	297	53Q
Firat et al. [5]	Evaluation of the treemap test and educational treemap tool	T	The Book of trees[34], students' designs, Google search	Classroom	Desktop Computers	25	57Q
User Studies with PCP							
Siirtola [31]	Testing the effect of linking RM and PCP	RM, PCP	RM-PCP Browser (C)	Controlled User-study	Desktop Computers	20	20T
Yang et al. [30]	Assessment of the IHD framework	SC, SG, SP, PCP	IHD Framework (C)	Controlled User-study	Desktop Computers	20	1T
Siirtola and Räihä [13]	Evaluation of the PCP vs SQL	PCP	PCP Explorer (C)	Controlled User-study	Desktop Computers	16	16Q
Lind et al. [35]	Evaluation of the new PCP axis layout	PCP	Many-to-Many Layout (C)	Controlled User-study	Desktop Computers	12	21Q
Claessen and Wijk [36]	Evaluation of the new PCP prototype usability	PCP	FlinaView (C)	Controlled User-study	Desktop Computers	10	7Q
Rosenbaum et al. [37]	Testing effectiveness of the new style in pattern detection	PCP	Progressive PCP (C)	Crowdsourcing	Amazon MTurk [32]	43	20Q
Palmas et al. [38]	Evaluation of the new vs classic PCP	PCP	Edge-bundling Layout (C)	Crowdsourcing	University email	137	2T
Kanjanabose et al. [39]	Comparing SPs and PCPs	PCP, SP	Manual	Controlled User-study	Desktop Computers	42	4T
Kwon and Lee [40]	Investigating the efficacy of the online learning environments	PCP	Manual	Crowdsourcing	Amazon MTurk [32]	120	18Q
Our Work	**Evaluation of PCP literacy barriers and test development**	**PCP**	**High-D [41], Mondrian [42], Quadrigram [43], PCP Tool, Xmdv [12], XDat [44]**	**Crowdsourcing**	**Amazon MTurk [32], Qualtrics [45]**	**60**	**28Q**

Table 2: An overview of the related literature of visualization literacy and PCPs with user-studies. The columns are: user-study themes, visual designs tested, image generation tools and methods, evaluation techniques, the tools used for evaluation, the number of participants included in the user-studies and number of tasks (**T**) or questions (**Q**) asked are provided. The evaluation technique that each research paper uses is categorized into: controlled user-study, classroom setting, and crowdsourcing. Abbreviations used for visual designs include **BC:** Bar Chart, **CM:** Choropleth Map, **H:** Histogram, **LC:** Line Chart, **PC:** Pie Chart, **RM** Reorderable Matrix, **PCP:** Parallel Coordinates, **P:** Pictographs, **S:** Spiral Chart, **SC:** Stack Chart, **SP:** Scatterplot, **T:** Treemap, **IHD:** Interactive Hierarchical Display. (**C**) indicates that the paper introduces a customized tool for image generation.

Siirtola and Räihä [13] discuss the methods of interaction found in PCP browsers and review them in accordance with existing user interface guidelines. An empirical study is performed in which the usability of PCPs is tested. They indicate the rich interaction opportunities of parallel coordinates and illustrate the value of interactivity for the method. We also incorporate an interaction to display the correspondence between CCP and PCP designs that advances user understanding.

Lind *et al.* [35] provide a new layout of axes for multiple PCP displays to assist users. The display of many-to-many relational parallel coordinates is designed to enhance the visual exploration of relations between variables and can be used to explore objects or object groups. Our work differs by presenting a standard PCP layout rather than a new, enhanced design.

The goal of Claessen and Wijk [36] is to allow users to freely identify and position flexible, linked coordinate axes and specify novel visual layouts by linking these axes with a flexible layout. This enables users to compare scatterplots, PCPs, and radar charts and also create a highly customized layout. The method is tested with 10 users who considered the idea easy to grasp and highly encouraging. This work presents new PCP features e.g. axes layouts that can potentially provide future work for our study. Their study assumes that participants already have a basic knowledge of PCPs whereas ours does not.

Rosenbaum *et al.* [37] implement positive PCP to overcome data collection and design challenges due to processing large volumes of data. A systematic study was performed with 43 participants to compare the usefulness of progressive PCP with regular PCP. The participants were asked to perform a variety of exercises, such as recognizing patterns in instances and looking for similarities in various stages in refinement. The findings show there was no major difference between the two methods in terms of accuracy. However, progressive PCP were slightly quicker for pattern detection and, on average, just 37% percent of the data was required to identify the patterns. Their study requires that participants start with a basic PCP literacy level whereas ours does not.

Palmas *et al.* [38] introduce an edge-bundling technique using density-based clustering for each dimension. It enables the clustered lines to be rendered using polygons, significantly reducing the rendering time. They also develop attribute relations with this technique to promote multidimensional clustering. A web-questionnaire with two tasks is given to compare the classic PCP against the new visual design in a user-study. The link to

the questionnaire is sent to computer science students and researchers at a local university and in total 137 respondents are analyzed. Their user study assumes basic PCP literacy before participation, whereas our does not.

Kanjanabose *et al.* [39] conduct an experiment including 42 participants to compare scatterplots (SP) and PCP and measure user performance in terms of accuracy and response time using four specific tasks. Three levels of task difficulty are given users on three representations as data table, SP and PCP. Similarly, we display the connection between a CCP and PCP and gauge participants' level of PCP literacy.

Kwon and Lee [40] focus on parallel coordinates to study the impacts of multimedia learning environments for teaching data visualization to non-expert users by examining the effects of active learning theory. To research the efficiency for data visualization education, 18 questions are given to 120 participants in an experiment based on tasks such as mapping between data points and visual elements, data distribution, comparison and similarities. In this research, focus on parallel coordinates to study the impacts of multimedia learning environments for teaching data visualization to non-expert users by examining the effects of active learning theory. To research the efficiency for data visualization education, 18 questions are given to 120 participants in an experiment based on tasks such as mapping between data points and visual elements, data distribution, comparison and similarities. In this research, the video tutorial and the static tutorial approaches are similar to ours, but the study does not include viewing correlation with the standard method (CCP) to accelerate learning, identify barriers to the comprehension of a PCP, and develop a literacy assessment test.are similar to ours, but the study does not include viewing correlation with the standard method (CCP) to accelerate learning, identify barriers to the comprehension of a PCP, and develop a literacy assessment test.

The difference between research on literacy and research involving user-studies is that, in general, user-studies on PCPs assume a basic prior knowledge of the visual design whereas literacy studies assume no prior knowledge. This is because user studies typically focus on a specific PCP design optimization over the standard layout. Our work assumes no prior knowledge and develops a tutorial to promote literacy. Table 2 summarizes how our work compares to previous related work. This study is the first one of its kind focusing on PCPs. Unlike previous work, it compares and links CCPs with PCPs. We also compare the visual design of PCPs with the most popular visual designs to inform and identify the specific barriers to PCP literacy.

We design a novel pedagogical tool and conduct a crowdsourced user-study to gather evidence and study barriers to PCP literacy. The focus of our research is specifically on PCP literacy as we provide analysis and guidance to address the barriers to reading, understanding, and interpreting PCPs. Together, our work is a unique combination of barrier identification, PCP tool evaluation, PCP literacy test development, and a user study on PCP literacy.

3. The Challenges of Interpreting PCPs

PCPs provide a visual solution to study the properties of multivariate and high dimensional data. A PCP focuses on a continuous vertical dimension when positioning points along axes compared to a Cartesian Coordinate Plot (CCP) which follows both a continuous vertical and horizontal plotting process. We identify at least seven barriers to PCP literacy based on both a review of survey papers on this topic [27, 28, 29], our experience in teaching PCPs in the classroom with traditional slides (for over 10 years), and then assessing the coursework. We identify the following barriers:

- **(S) Space** PCPs use an alternative layout of space when compared to CCPs and other popular designs. The most popular visual designs are based on two orthogonal axes whereas PCPs are based on repeated (typically 2-10) parallel axes [10]. Unfamiliarity with this use of axes can create a barrier when interpreting PCPs.

- **(M) Multivariate** An obstacle to parallel coordinates comprehension is the requirement understanding of multivariate ($n \geq 3$) or high-dimensional ($n \geq 5$) data attributes and their relationships [46].

- **(C) Correlation** Identifying the correlation and relationships between data dimensions, as well as knowing how to interpret the slope of the edges, is one of the obstacles to parallel coordinates interpretation. The slope of edges between axes can convey a correlation. This barrier also requires understanding the statistical terminology alone, i.e. correlation, outside the context of PCPs, which make it more complex.

- **(D) Distribution** One of the barriers to interpretation is based on the spread of edges over screen space. The more uneven the distribution of edges is, the more difficult it may be to follow polylines as they cross

Figure 3: (a) Data visualization types surveyed by Lee *et al.* [10] from three sources: K-12 curricula, data visualization authoring tools, and popular news outlets. (b) The 12 most popular visual designs that compose the VLAT. Images courtesy of Lee *et al.* [10]

and obstruct one another. Overplotting can result in higher visual complexity and occlusion [28].

- **(O) Order** Parallel coordinates rely on an axis layout order that specifies placement of axes in screen space. The location of the axes may create an obstacle to understanding the relationships between data dimensions that are not adjacent neighbors in a PCP [28].

- **(P) Pathtracing** Path (or polyline) tracing is a new task in PCPs which is not required in CCPs.

- **(E) Edges vs Points** In a PCP edges correspond to points in a CCP. This conversion may cause cognitive challenges.

We also identify two general visualization barriers with respect to labels and legends. However, we chose not to investigate these general concerns further, because lack of labels and legends means that the PCP design is incomplete. For example, it is difficult to identify data axes with no labels. Although we do not investigate them further, for reference they are:

- **Labels** Labels and minimum-maximum values on the axes facilitate user understanding. Missing axis labels can prevent PCP literacy.

- **Legend** When a color mapping is used in a PCP, a color legend presents the range of values of a data attribute. The absence of a color legend can obstruct understanding of a PCP.

Lee *et al.* [10] investigated the 12 most popular visual designs included in the education curriculum, and the most frequently used visual designs in news articles. From *their* Figure 3, we observe that parallel coordinates is *not* among the most popular graphical representations because a PCP is a visual design which is difficult to comprehend. Considering the identified barriers, PCPs have a number of characteristics that differentiate them from the other most common visual designs Lee *et al.* described:

- It uses a unique space and axis arrangement to plot points along vertical axes that represent data variables. If we look at Figure 3 carefully, all other designs use a Cartesian coordinate space except for pie charts and treemaps. Axes are generally orthogonal whereas in PCPs they are in parallel and there are often several of them (**Space**)

- Parallel coordinates design is used to display multivariate data. Figure 3 contains only one multivariate visualization (the stacked bar chart). We believe this is one of the barriers to a comprehension of a PCP. (**Multivariate**)

- The PCP is the only design that focuses on identifying direct relationships between multiple ($n \geq 3$) attributes. (**Correlation**)

- The PCP presents a high number of data records using polylines that cross parallel axes. (**Distribution**)

- The PCP is the only visual design that requires decisions with respect to the order of variates other than the treemap. (**Order**)

- Each data variate is associated with a specific category which may have different types of a attributes. (**Labels**)

It is infeasible to guarantee that these are all barriers with respect to interpret PCPs. There may be other barriers and investigating further barriers is a very good candidate for future work.

4. Developing a PCP Literacy Test

We propose a PCP literacy test to assess an individual's literacy skills with PCP images [15]. In order to develop an effective literacy test, we surveyed both related work and parallel coordinates tools for appropriate

Features & Tools	Color Mapped Polylines	Customizable Color map	C. Polyline Color Choice	C. Background Color	Axis Labels	C. Axis Labels	Min-Max Values	C. Min-Max Values	Read Text Data	Removable Axes
GGobi [47]		✓			✓				✓	✓
High-D [41]			✓		✓			✓	✓	✓
Mondrian [42]			✓		✓				✓	✓
Quadrigam [43]	✓	✓		✓	✓		✓	✓		✓
Sliver [48]		✓	✓		✓			✓		✓
PowerBI [49]	✓	✓		✓			✓		✓	✓
Spotfire [50]	✓	✓			✓	✓	✓	✓		✓
SPSS [51]	✓	✓			✓	✓	✓	✓		✓
XDat [44]		✓	✓	✓	✓		✓	✓	✓	✓
Xmdv [12]	✓	✓			✓	✓				✓

Table 3: Tools used to create PCP images for the literacy test. The table shows name the tool and some important features supported such as color mapped polylines, customizable color map, customizable polyline color choice, customizable background color, axis labels, customizable axis labels, min-max values, customizable min-max values, ability to read text data, and removable axes. A green cell indicates support for the corresponding feature.

software and related data while considering the various software features. Our primary hypothesis is that the students taught with the interactive tool demonstration will perform better on the literacy test than those taught with static slides. We start the process of creating appropriate PCPs for the literacy test by exploring PCP tools and datasets available online. Then we use these datasets and tools to create quality PCP images. After this stage, we choose the highest quality images to include in our literacy test. We develop questions with multiple-choice answers for each PCP image after the PCP image selection. We then compile 28 questions for the PCP literacy test to evaluate and compare users' PCP literacy levels in an experimental setting before and after watching an educational video tutorial (see Section 6).

4.1. Identifying PCP Tools

We reviewed visualization tools that create PCPs using online software collections provided by Keshif [54], Kirk [55], and Liu *et al.* [56] (see Table 3). In addition, we searched for the most frequently used visualization tools to generate PCPs using Google. We also examined students' submissions of an information visualization assignment for the Data Visualization course we taught at Swansea University. Students created parallel coordinates examples with a description of which tools or technologies were used and the datasets used to create their images (see Figure 4). We incorporate

Figure 4: Parallel coordinates tool selection from computer science students' on information visualization assignments in 2018 and 2019.

teaching experience to inform and identify barriers to PCP literacy. Over the 10+ years of assessing these assignments, we see some of the same errors repeatedly.

We assessed and compared each tool systematically based on the quality of the rendering, the use of color, the presence and legibility of axis labels, and clearly marked axis scales. By quality of rendering, we mean rendering without too much line aliasing. We do not have a specific color that we believe that is required, but generally darker colors on a light background with a good level of contrast. We looked for good quality color mapping too. Moreover, axis labels are necessary in order to interpret what each axis is for particular experiment. After testing PCP tools from a variety of sources, we identified a collection of tools to create quality PCPs for the literacy test as follows: **Ggobi** [47], **High-D** [41], **Mondrian** [42], **Quadrigam** [43], **Microsoft PowerBI** [49], **Sliver** [48], **Spotfire** [50], **IBM SPSS Statistics** [51], **XDat** [44], and **Xmdv** [12]. These tools provide quality parallel coordinate plots with clear designs, coloring, and labeling options (see Table 3 and Figure 5).

Table 3 indicates important features of the tested PCP tools. The color-mapped polylines column identifies the ability to apply a color map to a set of polylines while the customizable color map column indicates the option to

15

Dataset Name	Number of Records	Number of Dimensions	Name of Dimensions	Data Source	Description
Car	392	7	miles per gallon, cylinders, horsepower, weight, acceleration, year and origin	**Xmdv** [52]	Compares cars produced between 1968 and 1983.
Cereal	77	11	calories, protein, fat, sodium, fiber, carbohydrates, sugar, potassium, vitamins and minerals, display shelf, weight, and number of cups for each serving	**Xmdv** [52]	Provides information on nutritional properties in specific cereal products.
Chemical Elements	109	33	name, period group, chemical symbol, atomic mass, year of discovery, density, melting point, boiling point, enthalpy of fusion, enthalpy of vaporisation, molar entropy, enthalpy of atomisation, etc.	**High-D** [53]	Records characteristic attributes of chemical elements in the periodic table of elements.
Coal Disaster	191	5	months, years, day of year, interval, and deaths	**Xmdv** [52]	Records a number of coal-mining accidents between March 15, 1851 and March 22, 1962.
Energy	51	12	state, total energy consumption, per capita energy consumption, residential sector, commercial sector, industrial sector, transportation sector, petroleum, natural gas, coal power, hydroelectric power, and nuclear electric power	**Xmdv** [52]	Records the energy consumption in US states in terms of energy types and sectors that use the energy.
Olive oil	574	11	area, region, palmitic, palmitoleic, stearic, oleic, linoleic, linolenic, arachidic, eicosenoic, test/training	**Mondrian** [42]	Describes eight chemical measurements on different olive oil samples produced in different Italian regions.
Tipping	244	7	total bills, tip collected, gender, smoker, day, time (day/night), size of team worked	**Mondrian** [42]	The data was collected to show tipping behavior in a restaurant located in a shopping mall.
US Election	3111	53	name, state name, Bush, Kerry, Nader, total, male, female, obese, unemployed, rent, pcturban, urbrural, etc.	**Mondrian** [42]	Records information on the 2004 US presidential election that includes characteristics of on voters' demographics with state name for three candidates.
US Population	75	5	year, total population, percent change, resident population, and civilian population	**Xmdv** [52]	Provides US Population census data from 1900-2006 in thousands.
Venus	8784	7	date, hour, latitude, longitude, PV Plasma, Velocity (km/sec), PD Plasma Density (no/cc) and PT Plasma Temperature (K).	**Xmdv** [52]	The Venus atmospheric data is time-oriented and collected from a NASA mission.

Table 4: The original datasets used to develop the PCP literacy test (before modification). The table indicates a name of dataset, number of records, number of dimensions, name of dimensions, data source, and description. Column 5 provides the URL of each data set via citation.

change the color map. Customizable polyline color is important for changing the color of individually selected polylines. Customizable background indicates the option to update the background color. Axis labels refers to visible labels at the top or bottom of each axis and some tools facilitate modification of axis labels through the software (customizable axis labels). This feature is vital for creating legible PCP images to end users. We use customizable polyline color to create images of polylines with special colors that can be tracked across the display and hence multiple dimensions. We modified some of axes labels to make them more legible. For example, if the axis labels are too long or small, we shortened the labels or increased the font size. Min-max values indicates the software displays min and max attribute values and customizable min-max values implies that user can change these values using the software. As a consequence of changing min-max values, the axes should be re-scaled. We modified a number of min-max values to simplify some of the PCP images including removing a number of unnecessary decimal places. Some tools can only read numerical values. This is important because some datasets feature text data. Tools that can read text data are indicated under the *Read Text Data* column. The remove axes feature can enable making
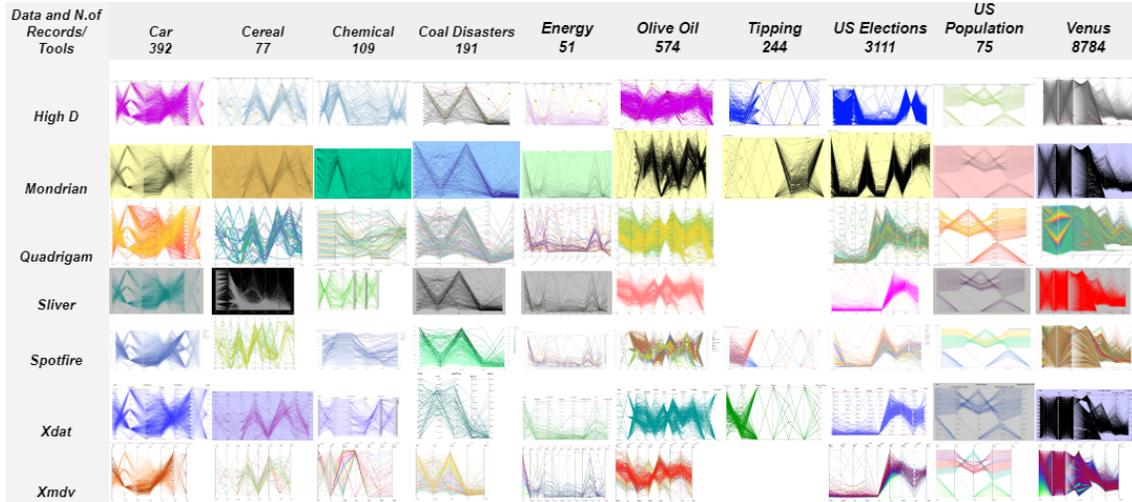
Figure 5: Parallel coordinates matrix used to develop the PCP literacy test. The matrix indicates a name of dataset and tools used to create PCP images. Since Tipping dataset has many text attributes, PCPs images were not created using tools that cannot read text data (see Table 3).

some axes visible or invisible on the image. We incorporated this feature to lower the dimensionality of some the datasets thus simplifying some of the images, for example, keeping the number of dimensions ≤ 10.

4.2. Exploring Datasets

Creating appropriate PCPs for the literacy test is important to assess users' literacy levels. Therefore, we carefully sought to find appropriate and interpretable multivariate datasets through the websites of tools (e.g. Xmdv datasets [52]) that we selected for creating PCP images. In order to select appropriate datasets for the literacy test, we constrained them using the following criteria:

- Multi-dimensional including 4 or more dimensions,

- A minimum size of 5 records and a maximum of 10,000,

- Studying a theme or behavior that is not too common,

- Clearly documented data dimensions.

The datasets search is intended to create an adequate variety of images and enable test takers with various literacy skill levels to attempt a range of questions. We found 10 datasets identified in Table 4. We noticed many previous related papers only test one or two data sets. We wanted a wider variety than many of the previous related work. We also created images with a balance of clusters, correlations, patterns, and outliers across images. The axes are generally ordered in the same order as the original data.

4.3. PCP Image Generation Process

Together, we used these datasets and tools to create quality PCP images for the literacy test questions. We imported the datasets using each tool, removed some redundant dimensions, updated some polylines, and background colors for higher quality images. We encountered some difficulties generating PCPs due to some limitations of tools. Not every tool provides the same features such as displaying visible and large axes labels, a color map, showing min-max values, and reading non-numerical values (see Table 3). These barriers inspired us to try each combination of dataset with different tools to determine the most appropriate images for the test (see Figure 5). Only some images in Figure 5 were selected – we tried to choose only those which were of high quality.

As color is useful for a range of parameters, such as aesthetics, color mapping, and visibility options, we used color to experiment with these parameters. For example, color mapping is used for a data variable on one dimension and sometimes. Color is also used to modify the background in order to display a contrasting polyline color. Moreover, we wanted to make sure that clusters, patterns, outliers, and correlations were balanced across images. Our image selection criteria was also based on using anti-aliasing and color mapping for polylines. We left out the images that have a distracting background or background color that interferes with the polyline color. For the images that ended up in the test, refer to Table 1.

4.4. Deriving a PCP Literacy Test

The test was created to examine the understanding of different aspects of parallel coordinates design and the methods of visual exploration and analysis. For the literacy test, a matrix of parallel coordinates images for each dataset and tool (see Figure 5) were used to inform the most appropriate examples which provide a good color choice and display data attributes correctly in an image. From the collection, we chose the highest quality images
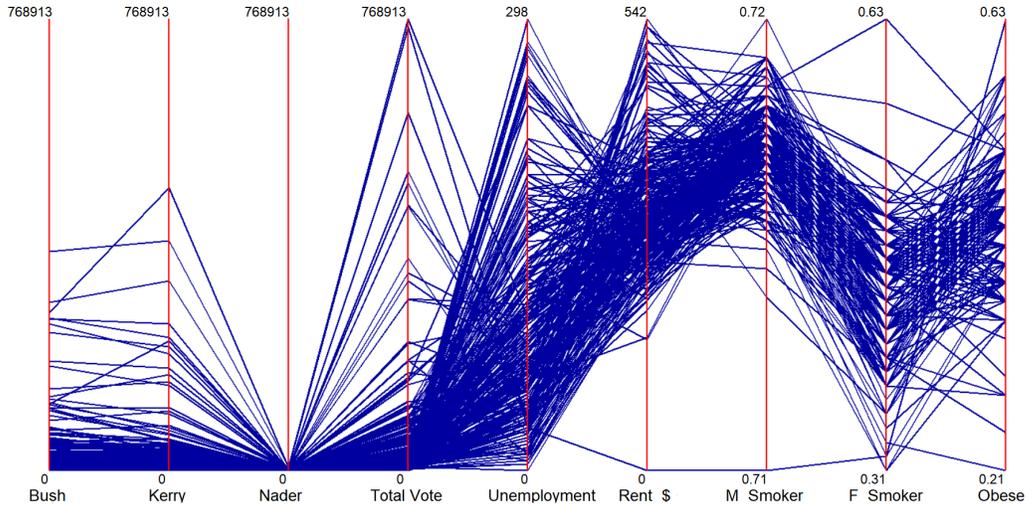
18

Figure 6: An example PCP literacy test question. *Which variable has an indirect correlation with the unemployment rate?* Options: *A) Bush, B) Kerry, C) Nader, D) Not sure, E) None of the above.* The full set of questions can be found at [15].

based on the properties in Table 3. Once the PCP image selection was completed for the test, we prepared questions with multiple-choice answers for each PCP image. Test questions were derived based on previous user-studies [36, 39, 40, 13], involving parallel coordinates and targets the barriers identified in Section 3. We chose to derive some test questions from these studies since they have been through a refereeing process and approved by reviewers.

PCP Question Classification: We categorized each question with respect to the barriers we identified in Section 3. An example is shown in see Figure 6.

- **(S) Space** requires the user to search the PCP for specific data points.

- **(M) Multi-variate** requires an understanding of two or more variables.

- **(C) Correlation** identifies questions that target understanding of the relationship between two attributes; direct, indirect, or no relationship.

- **(D) Distribution** indicates those questions that require an understanding of polyline distribution; characterized by how edges are distributed along an axis.

19

- **(O) Order** identifies questions that are influenced by axis order. (non-neighboring axes)

As a result, we compiled a total of 28 questions for the PCP literacy test. Each question in the test requires understanding at least two challenges to answer. For example, Figure 6 displays information from the 2004 US presidential election that includes records of voters' demographics including the names of three political candidates. The data variables are presented as state name, candidates Bush, Kerry, and Nader the total number of votes, rent in US Dollars, unemployment, and obesity rates. The question is *"Which variable has an indirect correlation with the unemployment rate?"* In order to answer this question, the user is required to understand **Space**, **Multivariate**, and **Distribution** aspects. Figures 10 and 11 record the classification of the each test question in terms of PCP literacy barriers and indicate which aspects of PCPs a user is required to understand in order to answer it. Each image in the literacy test is accompanied by a description of each data dimension. We believe if we do not describe the data dimensions, it would be very difficult to answer the questions without knowing what the dataset and data columns are. However, designing a PCP literacy test that requires no data descriptions could be good future work, even though it is challenging.

5. Developing an Educational PCP Literacy Tool

In addition to a PCP literacy test, we developed a customized, interactive pedagogical tool to enhance a non-expert user's PCP literacy skills. Testing many different existing PCP tools in conjunction with developing the PCP literacy test was very informative when developing our own pedagogical software. Our tool is intended to facilitate the interpretation and exploration of multivariate data as well as enable users to create and interpret PCPs interactively. The software features correspondence between CCP and PCP views and links the points in Cartesian space with polylines in parallel coordinates space (see *PCP Literacy Tool Demonstration Video* [14]). The pedagogical tool is developed as a desktop application and the design is inspired by Alfred Inselberg's software [26] that draws the correspondence between CCP and PCPs. This is the way that Inselberg himself taught PCPs after many years of using slides and we added new functionality to his software. In particular, we increased the number of dimensions, included glyphs, labels, and color mapping. Developing PCP software based on this approach is a novel aspect of our work.
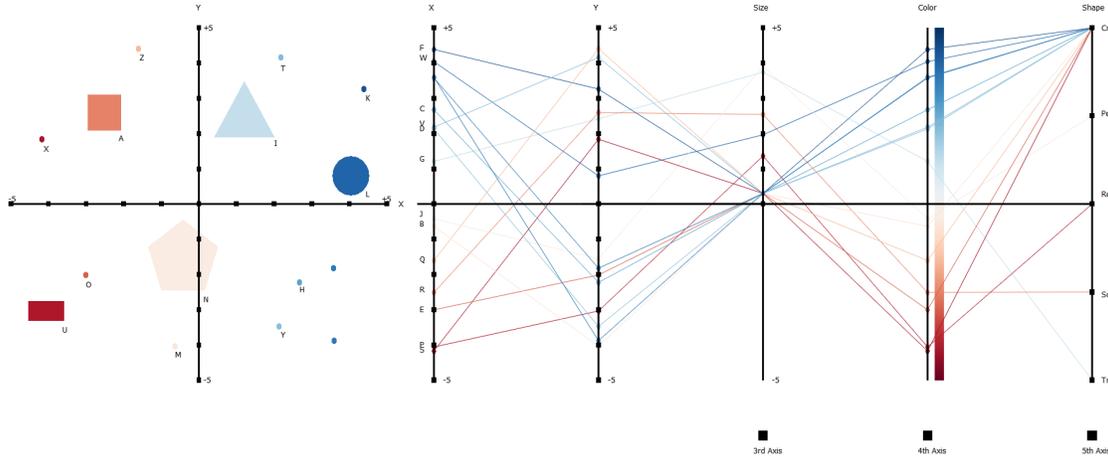
Figure 7: Pedagogical tool interface with Cartesian coordinate space (left) and the corresponding parallel coordinates plot (right). Our PCP Literacy Tool Demonstration Video can be found at [14].

Cartesian Coordinate Plot Features (CCP): Seeing the correspondence between the traditional CCP and the new PCP can help users understand the PCP. The interface (Figure 7, left) shows a CCP where users can specify points and interactively position them in a traditional 2D coordinate system by clicking on them. Moreover, a user can right-click on each point to resize, change the shape, return to the default shape/location, and remove the point. The color variable is mapped to the x-position of each point by default. Any update to a point will update the x-position and color accordingly.

Parallel Coordinates Plot (PCP) Features: The tool (Figure 7, right) presents a PCP with up to five attributes that represent the dimensions such as x-position and y-position in Cartesian space, the size of each point, color, and shape. Each component (edge or point) of the PCP corresponds to a component (point or edge) in the CCP. The size, color, and shape axes can be toggled interactively to observe the direct connection between dimensions in the CCP and PCP. For each point rendered in the CCP, an edge is created in parallel coordinates. The user can hover the mouse over any point or edge. This highlights and displays the correspondence between the CCP and PCP views. Interactive control of either point or edge updates the other in both

21

spaces. Moreover, users can create edges between x and y axes in the PCP that result in the corresponding points in Cartesian space with default size, color and shape settings. This feature facilitates the creation of points by drawing edges and increases the user's comprehension of the relationship between points and edges.

Menu and User Options: The menu options provide more features to the user. The file menu can save the points and corresponding polylines on the screen or load from a previous file as well as deleting all of the current points. These features were intended for a classroom-based experiment which was not possible. However, these features are useful for testing purposes and creating tutorial videos.

Labels: In order to convey the link between the points and polylines, each point and polyline can be labelled with a letter. This feature facilitates understanding by matching a point in the CCP view and corresponding polyline in the PCP with the assistance of the corresponding labels. Also, the user can toggle display labels and change the label font size.

6. Experimental Design and Crowdsource User-study

The purpose of this experiment is to evaluate the PCP literacy skills of participants with our novel test as well as seeking evidence of the educational PCP software efficacy compared to traditional slides in advancing novices' PCP literacy skills. Our hypothesis is that our tool will improve PCP literacy compared to traditional slides.

6.1. Experimental Setting

Our original intention was to conduct an in-class user-study similar to that of Firat *et al.* [5] to evaluate our pedagogical software tool. Hence, we believe our software helps both students and the general user. However, enabling the interaction of participants with the PCP software tool during an in-person lecture was no longer feasible. As the researchers were not teaching a visualization class at the time of the pandemic, it wasn't possible to conduct this study with the students at the university where the researchers are based. Thus, we chose crowdsourcing for an empirical evaluation of PCP images and software during this challenging period and we designed a crowdsourced user-study as opposed to a classroom-based one. This choice is inspired by the many crowdsourced experiments performed in the visualization literature e.g. Borgo *et al.* [57]. Crowdsourcing enables us to reach users from a

large and diverse pool of backgrounds and to evaluate both varying levels of PCP literacy and the efficacy of the educational PCP tool. We designed the experiment incorporating the checklist provided by Borgo *et. al* [57] for completeness and minimizing error. We split our PCP literacy test into two parts: pre-tutorial and post-tutorial for the study. Both tests contain 14 randomly assigned literacy questions from a larger collection. Each question consists of a PCP image, a description of each image and a multiple-choice question with a single correct answer. Each question is accompanied by a description of each data dimension. We believe if we do not describe the data dimensions, it would be very difficult to answer the questions without knowing what the dataset and data columns are. However, designing a PCP literacy test that requires no data descriptions can be a good future work even though it is challenging. We also used additional PCP questions as simple screening tests. Each PCP design, dataset, and question provided in the test varies in complexity. The tests [15] were administered using Qualtrics [45], an online survey tool for collecting data.

6.2. Condition Groups

The user-study participants were divided into two experimental conditions: SLIDES and SOFTWARE. We investigated the impact of our educational PCP software that shows the correspondence between a CCP and PCP, based on improving the PCP literacy of users by comparing the novel software to the traditional PCP slides. For the purpose of this study, we prepared a video tutorial for the SLIDES condition that explains the features of PCP with traditional lecture-style slides only [17] because this is the most commonly used presentation material in the classroom. Time constraints prevent the use of interaction with the dozens of visual designs taught in a visualization class. For the SOFTWARE condition, we prepared a video tutorial that briefly introduces PCPs with slides but also displays the interaction featured in the pedagogical tool demo by showing each of the features described in Section 5 [18].

Interaction– A hands-on approach to interacting with the educational tool is not part of our assessment because it introduces a new variable into the experiment (e.g. hardware). This would make the experiment more complicated and could confound the results. For example, the interaction is different depending on the input device being used (mouse, touchpad, trackball, touchscreen, joystick, etc.). In a crowdsourced experiment, we do not have as much control over the hardware input-device used. Moreover,

users often encounter static PCPs rather than tools with which interaction is possible. This is especially true in the classroom and in presentations in general. This also indicates the importance of studying every challenge with respect to PCPs that may enable or prevent users to advance PCP literacy. In some cases the investigation provided by interaction eliminates some barriers to PCP comprehension. Given the importance of interaction, studying this topic is a good future work direction (more information is provided in Section 8.3).

The SLIDES [17] and SOFTWARE [18] tutorial videos are approximately 10 minutes long. Overall, it took approximately 40 minutes for each participant to complete the study. We also believe that length of the tutorial video and the test impact the number of valid responses collected in the user-study. To mitigate for the bias between two groups in terms of the content taught, we delivered the same essential information about PCPs in both tutorial videos e.g. example PCP designs, relationships between axes (direct, indirect, or none). This means that with respect to the topics presented, the slides were the same for both groups. The essential difference between the SLIDES and SOFTWARE groups was the software demonstration video, which therefore comes down to the animation used in the software that interactively displays the links between a CCP and a PCP. Thus, we rely on participants previous knowledge of CCPs in the software. This is based on the assumption that most users have been taught CCPs in their general education. Their are likely a small number of exceptions. Both groups answer the same set of PCP literacy test questions, e.g. the independent variable.

6.3. Participants Screening

We recruited participants through Amazon Mechanical Turk (MTurk) and established a strict screening process to improve data quality. Participants were required to have a total of 1,000 or more approved HITs and a 95% or greater HIT approval rate on MTurk. Before the experiment started, we ruled out participants who self-reported that they were using a mobile phone since the PCP images require significant screen space due to their high level of detail. We also screened out participants who did not recognize the numbers in color blind test images as our PCP images do not use color-blind safe colors. We also identified participants that attempted take part of the study more than once. We used longitude and latitude data and looked at time stamp information to screen these participants. The responses obtained from those participants were removed from the data to analyze.

6.4. Pilot Study-1

Our experiment was initially started with a pilot study to test the experimental design and procedures. A total of 15 participants were recruited regardless of their background and all of them completed the tests. Although participants are informed on the importance of watching the video tutorial in the instruction video, our screening identified that only 3 participants spent the time necessary (or close) to watch the entire tutorial according the data provided by Qualtrics. This initial trial informed us that we need to ask questions specifically about the video tutorials to encourage participants to watch the tutorial video until the end and to pay attention to the tutorial.

6.5. Pilot Study-2

Based on the experience obtained from the first pilot study, we prepared 3 questions on the tutorial video with multiple correct answers (e.g., *Which topics were mentioned in the video tutorial?, What was the example data about on the last slide of the video tutorial?, Which one of the following was a data attribute in the last example?*) to assess one's engagement with the presented material. We followed the same procedure in the second pilot study with 10 participants. As a result, even if some participants watched the tutorial video in its entirety, all but 2 participants selected random answers to the video tutorial on all 3 screening questions. Consequently, we decided to carefully screen for attentive participants to gather quality data for the main study.

6.6. Experimental Procedure

We published the link of the survey on the MTurk website [32] and asked for Turkers' consent to participate in the study. With their agreement, participants were provided with an instruction video that describes the experimental procedure, followed by some demographics questions. After that, they began the pre-tutorial test which consists of 14 randomly selected PCP images and multiple-choice questions with the goal of determining the current level of a participant's PCP literacy. Upon the completion of a pre-tutorial test, one of the video tutorials were randomly assigned to participants (SLIDES or SOFTWARE). The participants were instructed to watch the video tutorial carefully until the end.

After the video tutorial, participants answered questions about the tutorial and took the post-video tutorial test. The correct answers to pre- and post-video tutorial tests were counted and participants were awarded 1 point

in both tests for each correct answer and 0 points for any incorrect answers. To evaluate a participant's PCP literacy using this test, we then calculated the percentages of correctly answered questions by adding all points up and dividing the result by the overall number of questions. These percentages were then used as the data samples for our analysis. After both tests, 7 open-ended questions were given to collect feedback about the experiment and measure their confidence in the material. Additionally, 4 questions are asked specifically about the educational tool to the participants included in the SOFTWARE condition (see Appendix). After giving the questionnaire for feedback collection, participants were provided a unique completion code by Qualtrics that is submitted to the survey page on the MTurk website to complete the study.

Each part of the PCP literacy test (pre- and post-tutorial video) begins with two multiple-choice PCP questions as screening checks before randomly selected PCP literacy questions (e.g., *How many parallel axes are there in the image?*, *What is the dataset about?*). Three questions about the tutorial video follow the completion of the video tutorial session. Correct responses to these simple questions about the videos, and the recorded time spent on playing the tutorial video, helped us screen for attentive user-study participants. Inattentive participants were excluded from the data.

6.7. Data Collection and Filtering

We continued with the recruitment process until we obtained 60 valid responses from the same number of participants for both experimental conditions in order to provide a balanced comparison and effective quantitative analysis. Initially, we identified a total of 202 attempts to participate in the experiment. After filtering the participants based on color vision deficiency, using mobile phones, or those who tried to participate more than once from the same location, 170 participants remained. In the next step, we looked at the time participants watched the video tutorial, the answers given by the participants to the screening questions, and the simple questions about the video tutorial. At the end of filtering process, we collected responses from a total of 60 participants which is comparable sample size to previous crowdsourced studies [57].

The number of participants involved in each crowdsourced user-study in Table 2 shows that three studies use fewer than 60 participants and three others use more than 60 participants. This puts our 60 participants at the median of the related literature. We also believe that the duration of the
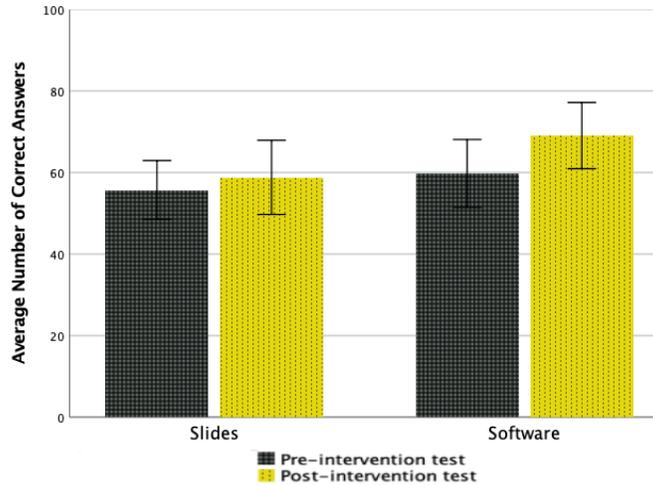
Figure 8: The percentage of correctly answered questions in the pre- and post-intervention tests for SOFTWARE demonstration and SLIDES groups. Error Bars (95% CI).

study has an impact on collecting data provided by attentive participants (approximately 40 minutes). Thus, we screened out a total of 110 participants (65%) from the experiment after about a week spent for collecting data.

Among the 60 participants (30 participants per condition SLIDES and SOFTWARE), the majority of the participants reported that they were from the United States (29 US, 15 India, 10 Europe, 3 Brazil, 1 Bangladesh, 1 Indonesia, 1 Pakistan). Participants (36 male and 24 female) had different degree levels: 28 Bachelor's, 17 Master's, 8 Ph.D., and 7 High School. The age of participants ranged from 18 to 60 with an average age of 33.4 years.

7. Results

Our primary hypothesis was that the students taught with the interactive tool demonstration would perform better on the literacy test than those taught with static slides. Our second hypothesis is that out PCP literacy test can be used to evaluate and improve PCP literacy.

7.1. Quantitative Results of Test Data: Slides vs Software Conditions

The percentages calculated for each participant were normally distributed, as indicated by the Shapiro-Wilk test for both pre- and post-intervention test groups. Hence, we used one-way ANOVA for our data analysis (significance level at $\alpha = 0.05$).

27

The pre-intervention test results did not differ significantly between the two groups: $F(1,58) = 0.564$, p = 0.456, $\eta_p^2 = 0.010$. Those participants who then watched the SLIDES video correctly answered on average 55.71% of the pre-intervention test questions (SD = 19.35%), and the participants who were in the SOFTWARE condition answered 59.76% of these questions (SD = 22.31%).

As for the post-intervention test results, those who were in the SLIDES condition on average responded correctly to 58.81% of the questions (SD = 24.36%), while the subjects who watched the SOFTWARE video answered 69.05% of the questions correctly (SD = 21.74%). This difference, however, was not significant: $F(1, 58) = 2.950$, $p > 0.05$.

The SLIDES group have seen a 3% increase in the percentage of correctly answered questions from pre-intervention test to post-intervention test. This increase was not significant: $F(1,29) = 1.796$, p = 0.191, $\eta_p^2 = 0.058$. On the other hand, the SOFTWARE group performed significantly better in the post-intervention test than in the pre-intervention test, having improved their results on average by 9%: $F(1,29) = 8.092$, p = 0.008, $\eta_p^2 = 0.218$ (Figure 8).

Further to these findings, we hypothesized that the software tutorial video would provide additional support in overcoming different barriers to understanding PCPs. Thus, we have also looked at the participants' performance in the different parts of the test aimed at measuring one's comprehension of different attributes of parallel coordinates. We did so by looking at the question classification based on the parallel coordinates features that could influence the participants' answers in the test.

In the post-intervention test, there was no difference in the results obtained by participants in both groups for the questions about Correlation (**C**), Distribution (**D**), and Order (**O**). However, participants who were in the SOFTWARE condition performed much better than participants in the SLIDES condition when answering the questions related to the Space (**S**) ($F(1,58) = 4.316$, p = 0.042, $\eta_p^2 = 0.069$) and Multi-variate (**M**) ($F(1,58) = 4.087$, p = 0.048, $\eta_p^2 = 0.066$) categories.

Within condition, comparing pre- and post-intervention, participants in the SLIDES group did not improve their performance significantly in any specific category of the questions. Within the SOFTWARE condition, our participants significantly increased the number of correct responses for the Multi-variate (**M**) category of questions (Table 5), but no significant improvement was observed in the other four question categories. The quantitative results are not as strong as we expected. However, we encountered improvements in

	SLIDES					SOFTWARE				
	Pre-Intervention	Post-Intervention	$F(1,29)$	p	η_p^2	Pre-Intervention	Post-Intervention	$F(1,29)$	p	η_p^2
S: Space	56.33 ± 25.26	59.57 ± 26.34	0.392	0.536	0.013	64.47 ± 27.73	72.70 ± 22.47	3.724	0.063	0.114
M: Multi-variate	52.90 ± 19.30	54.77 ± 27.72	0.122	0.729	0.004	55.93 ± 22.07	68.30 ± 24.00	10.153	0.003*	0.259
C: Correlation	47.60 ± 30.70	56.43 ± 31.88	1.018	0.321	0.034	46.00 ± 30.40	56.93 ± 29.71	3.602	0.068	0.110
D: Distribution	60.07 ± 31.34	63.97 ± 19.64	0.346	0.561	0.012	65.37 ± 30.32	72.90 ± 20.88	2.018	0.166	0.065
O: Order	42.50 ± 34.45	54.97 ± 33.37	1.778	0.193	0.058	41.03 ± 33.50	54.03 ± 32.06	3.862	0.059	0.118

Table 5: The results of pre- and post-intervention tests for the SLIDES and SOFTWARE groups ($M \pm SD$ in percentages), based on the categories of questions. Significant results are shown as follows: $*p < 0.05$.

comparing the pre-and post-intervention performance of the software group. There is also positive feedback in qualitative analysis in Section 7.2. In future work, we intend to undertake an in-class study to support our hypothesis as was originally intended.

7.2. Qualitative Analysis of Feedback Data

To analyze the feedback collected from our participants at the end of the survey, we followed a lightweight approach related to the initial inductive thematic analysis steps as described by Braun and Clarke [58]. We divide the themes based upon whether they come from questions about the PCP literacy test, or from questions about our PCP literacy tool.

PCP Literacy Test –

We identified three themes about the PCP literacy test questions: *Impact of the tutorial video*, *Distribution*, and *Correlation*.

Impact of the tutorial video: Participants were asked to reflect on their experience of answering post-intervention test questions. The feedback from some participants in both conditions indicates a high difficulty level of the test questions although they watched the tutorial video and answered randomly chosen questions. Some 14 participants stated questions were generally difficult: *"All questions are very difficult to find the answers."* (P1) and *"More difficult but not much"* (P15). This result reveals that although the test questions have varying difficulties, the PCP literacy skills of these participants are weak. However, 16 participants (27%) stated that the post-tutorial test was easy or became easier after watching the video tutorial: *"After the tutorial, it is easy to identify what is meant by the image, the only normal difficulty of evaluating a problem"* (P42). It is clear that most of participants found the video tutorial very helpful in Figure 9. Participants also indicated an increase in confidence after watching the tutorial video.
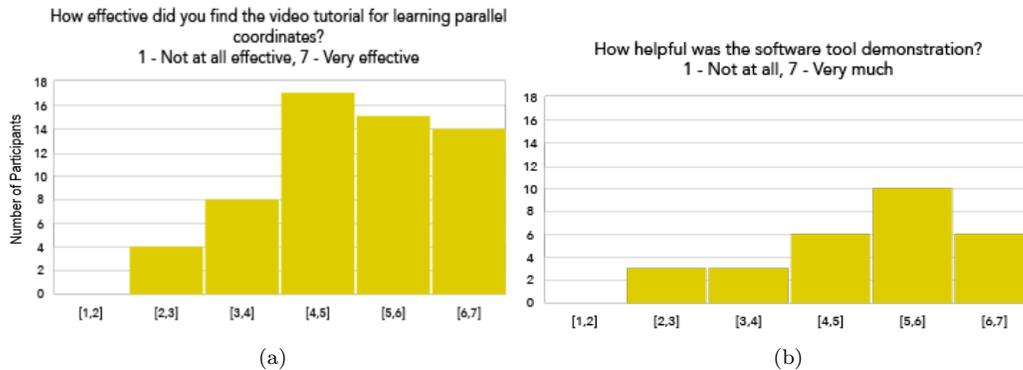
Figure 9: Participants' answers to feedback questionnaire about video tutorials.

Distribution (D): A high number of data records using polylines that cross parallel axes result in overplotted PCP designs – a barrier to PCP literacy. This challenge may be the result of not understanding crowded PCP images. This can play a significant role in participants' performance during the test. This difficulty is expressed by 5 participants (8%): *"There were complex plots with intermeshing lines being hard to follow the destination"* (P33) and *"[The questions were] very difficult because the data showed was very messy"* (P48).

Correlation (C): Observing values and axes labels to understand the relationship between two neighboring data attributes (direct, indirect, or none) is required to answer correlation questions correctly. Consistent with the quantitative data analysis, the qualitative feedback collected from participants in both groups did not show statistical improvement in every category (except the Multi-variate and Space categories in the SOFTWARE group). According to the feedback data, some 5 participants (8%) stated that they had difficulty understanding the correlation topic and answering the questions related to it: *"Before and after [the tutorial video], I still had the most difficulty with direct and indirect correlation."* (P23), *"I was torn between the direct and indirect."* (P27), and *"I had a hard time remembering the difference between direct and indirect."* (P47). This study gathers both qualitative and quantitative evidence identifying correlation as a barrier to PCP literacy. We hypothesize that this is due to the mathematics background required to understand this topic.

Educational PCP Software – We also evaluated the feedback directly relevant to our pedagogical application. The feedback was collected from

participants who were provided the tutorial video of the educational PCP software. Most of the participants (73%) reported that the educational PCP tool was effective and helpful for a better understanding of the relationships between data attributes. We coded the feedback based on the features of the software tool that were recognized as having a positive learning effect on the participants. We identified two aspects: *Multi-variate*, and *Correspondence and Animation*.

Multi-variate (M): Some six participants (20%) identified that the PCP tool was helpful to understand the PCP concept and convey the mapping of multivariate data which is one of the main challenges of high dimensional data. Participants stated: *"Connecting multiple data [attributes] in a single graph is better to get information easily."* (P41) and *"It [the tool] helps in understanding data with more than 2 variables."* (P43), and *"It [the tool] provides a direct visual demonstration of the mapping between axes."* (P50).

Correspondence and Animation: The educational tool demonstrates the correspondence between views of the CCP and the PCP using animation. Some five participants (17%) stated how the feature enabled them to understand the data through interaction between plots: *"It [the tool] provides to represent the data on both charts at the same time that explained everything clearly."* (P33) and *"It gave a good real-time illustration of the relation between the two coordinate types, and it [the tool] gave a clear way of displaying the data, with a more easily identifiable scale."* (P38), and *"Being able to see the changes from parallel coordinates chart to Cartesian chart made for better understanding of the relationships between data attributes."* (P56)

The Summary of Evidence– The results of quantitative analysis indicate that the software group has experienced a 9% increase in the percentage of correctly answered questions from the pre-intervention test to the post-intervention test while the slides group had improved their results on average by only 3%. According to the feedback collected from our participants at the end of the survey, most of the participants (73%) in the group who watched the software video reported that the educational PCP tool was effective and helpful for a better understanding of the relationships between data attributes. They rated the software tool demonstration average of 5-6 out of 7 (1-Not at all, 7-Very much) in terms of being helpful (see Figure 9b) and responded our question on the efficiency of the software tool positively e.g. *"Because it [the tool] allows learners to see how the graph works in action."* (P123) and *"It [the tool] gave a thorough explanation. And so, I felt like I could move my way around more effectively."* (P105). These findings
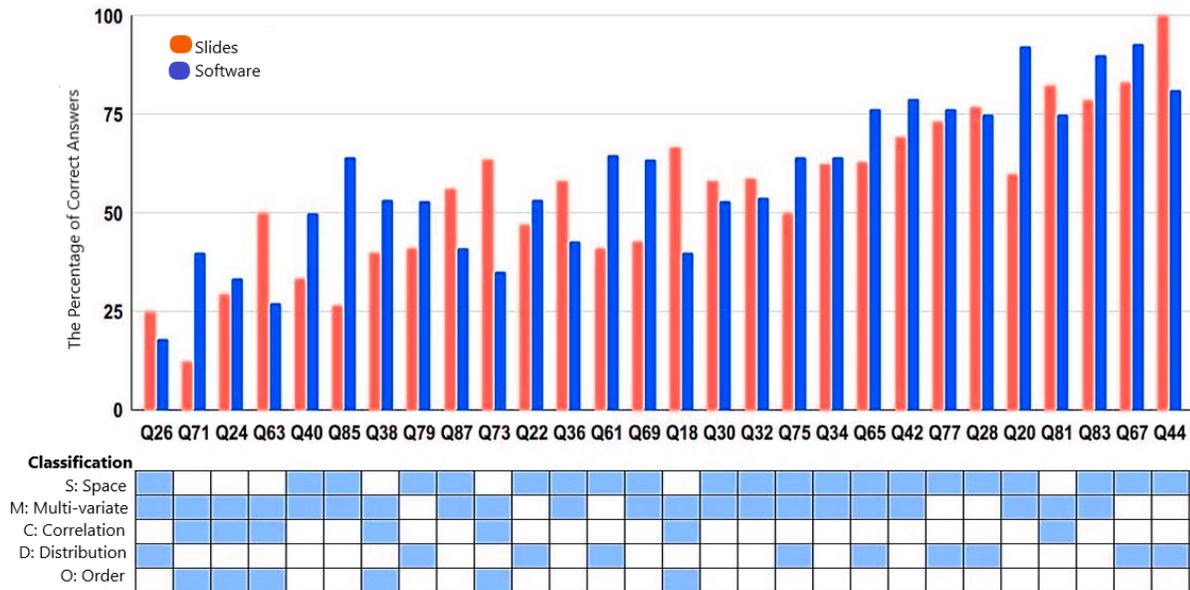
Figure 10: The percentage of correct answers on the pre-tutorial test by groups and the classification of questions. The questions are ranked from the most difficult to the easiest.

provide evidence that the software tool facilitates PCP literacy and supports our hypothesis.

8. Discussion

8.1. Designing Questions for the PCP Literacy Test

Figures 10 and 11 couples the classification of PCP literacy test questions with percentage of correct answers given by the participants in both pre- and post-tutorial tests. Questions are ranked from the easiest to the most difficult based on the percentage of correctly answered questions for both groups. The corresponding testing categories of each question are given below (see Appendix for example questions). Originally, we thought we could design test questions that isolate the different cognitive processes required to interpret the PCP images. However, this turns out not to be true. Observers usually have to understand two or three cognitive properties of PCPs in order to successfully answer PCP questions. We hypothesize that this confounding effect is itself a PCP literacy barrier. In addition, the results indicate that some participants may not be very familiar with the new terminology. Understanding the terms may also be a barrier to PCP literacy (e.g. Correlation).
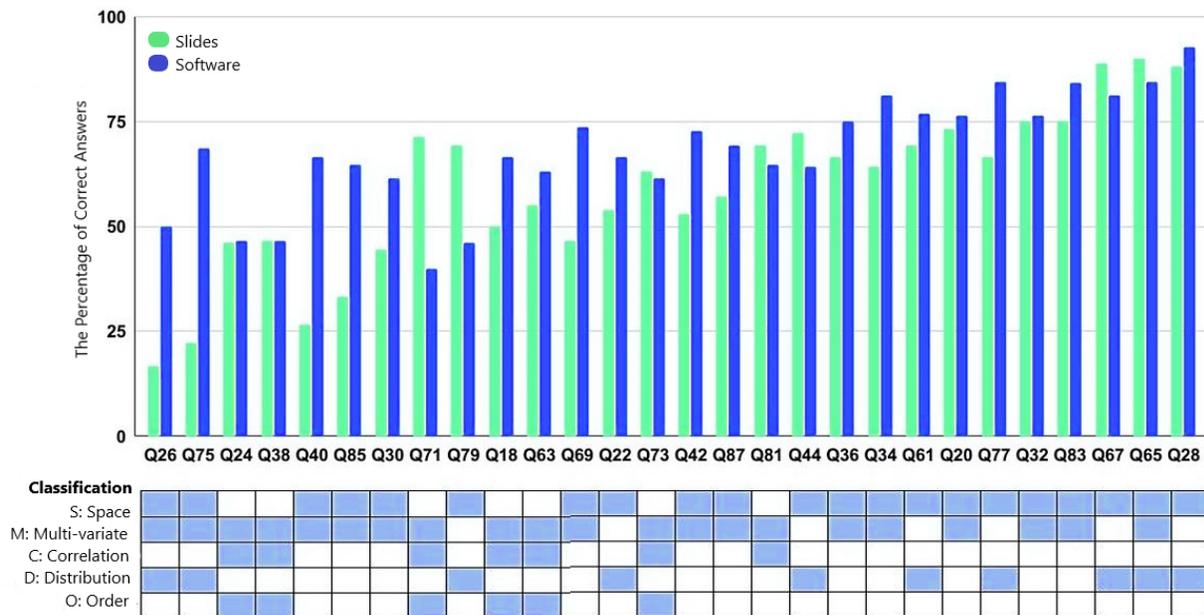
Figure 11: The percentage of correct answers on the post-tutorial test by groups and the classification of questions. The questions are ranked from the most difficult to the easiest (different to pre-tutorial).

One of the participants stated that *"There were a lot of terminologies to remember, and so that increased difficulty."* (P44). This indicates reinforcing the prerequisite terminology is an essential step in advancing a non-expert users' PCP literacy level.

8.2. The effectiveness of the PCP Literacy Tool

The pedagogical PCP tool shows the correspondence between CCP and PCP views to improve PCP understanding. The two views are connected and synchronized such that the mapping of multivariate data to geometric primitives is clearly illustrated. To lessen the impact of barriers to PCP literacy, the PCP tool facilities toggling axes in the display between different data attributes **(M, C, O)**. Points in the CCP can be added, deleted, or moved to demonstrate data distribution **(D, S)**. The experimental results and Figure 9b reveal the positive effect of the educational tool video tutorial on PCP literacy, based on participants' experience with the PCP tool demonstration. Although, participants did not explicitly address all individual barriers to

understanding PCPs in their feedback, overall feedback on the tool was very positive as indicated in Figures 9 and A.12.

Answering the literacy test questions requires perceiving more than one property of PCP. To help others who are not significantly influenced by the literacy tool, we can add new features such as loading new datasets and axis reordering axes as future work. A user-study dedicated to interaction in a controlled lab (an experiment which is currently not possible) might also be insightful.

8.3. Limitations and Future Work

Distance learning, a growing trend, offers a solution to access lecture material online as well as facing disadvantages such as difficulty staying engaged or receiving immediate feedback, etc. Our crowdsourcing user study features similarities to distance learning settings in which a video tutorial is provided to instruct the PCP design. Our findings reveal only 35% of the participants (60 out of 170) spent the time necessary (or close) to watching the entire tutorial. This indicates that keeping students engaged while watching a tutorial video online is a non-trivial challenge. We were not anticipating this barrier. For future research, we can conduct a similar experiment in a classroom setting and offer participants interaction with the educational PCP software instead of a video tutorial. Thus, we can examine the results to investigate improvement between crowdsourcing and classroom-based user-study settings.

Moreover, further studies are recommended to reinforce the effectiveness of the educational tool where interaction is necessary. Also, we intend to teach PCP design to novices by presenting the correspondence between a CCP and PCP, including advanced PCP features (e.g. axis flipping, axes reordering, removing axes) in the pedagogical PCP tool. Improving the PCP literacy test with a wider variety of datasets and asking users to provide their observations about a PCP design using a think-aloud protocol would be interesting for further study.

9. Conclusion

In this paper, we investigate barriers to PCP literacy based on the research literature and our teaching experience in data visualization classes. We provide insight into the barriers to a complete understanding of the concept of the PCP. A novel literacy test is presented to evaluate non-expert

user's literacy skills with a range of PCP images using a collection of tools and datasets. We also collated the most frequently used visualization tools and multivariate datasets used to generate PCPs for our literacy test.

Moreover, we developed an educational PCP tool to facilitate the interpretation and exploration of multivariate data as well as enabling users to learn how to create and interpret PCPs interactively. The software features correspondence between CCP and PCPs views and connects the points in Cartesian space with polylines in parallel coordinates space. Results of the user study reveal that participants taught with the pedagogical application tutorial video performed better on the PCP literacy test. Participants indicate that the illustration of the relation between the two views offers a simple way of improving PCP literacy as well as understand data with more than two variables.

References

[1] S. Bateman, R. L. Mandryk, C. Gutwin, A. Genest, D. McDine, C. Brooks, Useful junk? the effects of visual embellishment on comprehension and memorability of charts, in: Proceedings of the SIGCHI conference on human factors in computing systems, 2010, pp. 2573–2582.

[2] E. M. Peck, S. E. Ayuso, O. El-Etr, Data is personal: Attitudes and perceptions of data visualization in rural pennsylvania, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–12.

[3] J. Boy, R. A. Rensink, E. Bertini, J.-D. Fekete, A principled way of assessing visualization literacy, IEEE Transactions on Visualization and Computer Graphics 20 (12) (2014) 1963–1972. `doi:10.1109/TVCG.2014.2346984`.

[4] B. Alper, N. H. Riche, F. Chevalier, J. Boy, M. Sezgin, Visualization literacy at elementary school, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM, 2017, pp. 5485–5497. `doi:10.1145/3025453.3025877`.

[5] E. E. Firat, A. Denisova, R. S. Laramee, Treemap Literacy: A Classroom-Based Investigation, in: Eurographics 2020 - Education Papers, The Eurographics Association, 2020, pp. 29–38. `doi:10.2312/eged.20201032`.

[6] K. Börner, Data visualization literacy, Keynote Talk, IEEE Transactions on Visualization and Computer Graphics (2019).

[7] M. Romero, M. Velez, G. McInemy, D. Silver, M. Chen, Towards visualization literacy, EuroVis Workshop (2014).
URL https://www.kth.se/profile/marior/page/eurovis-2014-workshop-towards-visualiza

[8] S.-H. Kim, J. Boy, S. Lee, J. S. Yi, N. Elmqvist, Towards an open visualization literacy testing platform, IEEEVIS 2014 Workshop (2014).
URL http://visualizationliteracy.org

[9] K. Börner, A. Maltese, R. N. Balliet, J. Heimlich, Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors, Information Visualization 15 (3) (2016) 198–213. doi:10.1177/1473871615594652.

[10] S. Lee, S.-H. Kim, B. C. Kwon, VLAT: Development of a visualization literacy assessment test, IEEE Transactions on Visualization and Computer Graphics 23 (1) (2017) 551–560. doi:10.1109/TVCG.2016.2598920.

[11] R. Kosara, F. Bendix, H. Hauser, Parallel sets: Interactive exploration and visual analysis of categorical data, IEEE transactions on visualization and computer graphics 12 (4) (2006) 558–568. doi:10.1109/TVCG.2006.76.

[12] Xmdv Tool, http://davis.wpi.edu/xmdv/, Last accessed: (2021).

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

[14] A Tool to Support Parallel Coordinates Literacy Video, https://vimeo.com/456884883, Last accessed: (2020).

[15] Parallel Coordinates Literacy Test, https://swanseasom.au1.qualtrics.com/jfe/form/SV_2gxKduwQjtAQIuN, Last accessed: (2020).

[16] Instructions for the Experiment Video, https://www.youtube.com/watch?v=ReNKoU9AI5U, Last accessed: (2020).

[17] Introduction to Parallel Coordinates: A Tutorial (Using Slides), `https://www.youtube.com/watch?v=_G7eJHtgH5Y`, Last accessed: (2020).

[18] Introduction to Parallel Coordinates: A Tutorial (Using Software), `https://www.youtube.com/watch?v=JUfTNNlv97A`, Last accessed: (2020).

[19] L. McNabb, R. S. Laramee, Survey of surveys (SoS)-Mapping the Landscape of Survey Papers in Information Visualization, in: Computer Graphics Forum, Vol. 36, Wiley Online Library, 2017, pp. 589–617. `doi:10.1111/cgf.13212`.

[20] E. E. Firat, A. Joshi, R. S. Laramee, Interactive visualization literacy: The state-of-the-art, Information Visualization (2022).

[21] E. E. Firat, R. S. Laramee, Towards a survey of interactive visualization for education, in: Proceedings of the Conference on Computer Graphics & Visual Computing (CGVC), Eurographics Association, Swansea, UK, 2018, pp. 91–101. `doi:10.2312/cgvc.20181211`.

[22] D. Rees, R. S. Laramee, A survey of information visualization books, Computer Graphics Forum 38 (1) (2019) 610–646. `doi:10.1111/cgf.13595`.

[23] D3.js: Data-Driven Documents, `http://d3js.org`, Last accessed: (2020).

[24] P. Ruchikachorn, K. Mueller, Learning visualizations by analogy: Promoting visual literacy through visualization morphing, IEEE Transactions on Visualization and Computer Graphics 21 (9) (2015) 1028–1044. `doi:10.1109/TVCG.2015.2413786`.

[25] M. D. Reckase, Multidimensional item response theory models, in: Multidimensional item response theory, Springer, 2009, pp. 79–112.

[26] A. Inselberg, Parallel coordinates, Springer, 2009.

[27] A. Dasgupta, M. Chen, R. Kosara, Conceptualizing visual uncertainty in parallel coordinates, in: Computer Graphics Forum, Vol. 31, 3pt2, Wiley Online Library, 2012, pp. 1015–1024. `doi:10.1111/j.1467-8659.2012.03094.x`.

[28] J. Heinrich, D. Weiskopf, State of the art of parallel coordinates, in: Eurographics State-of-Art reports (STARs), Eurographics Association, 2013, pp. 95–116. `doi:10.2312/conf/EG2013/stars/095-116`.

[29] J. Johansson, C. Forsell, Evaluation of parallel coordinates: Overview, categorization and guidelines for future research, IEEE Transactions on Visualization and Computer Graphics 22 (1) (2015) 579–588. `doi:10.1109/TVCG.2015.2466992`.

[30] J. Yang, M. O. Ward, E. A. Rundensteiner, Interactive hierarchical displays: a general framework for visualization and exploration of large multivariate data sets, Computers & Graphics 27 (2) (2003) 265–283.

[31] H. Siirtola, Combining parallel coordinates with the reorderable matrix, in: Proceedings International Conference on Coordinated and Multiple Views in Exploratory Visualization-CMV 2003-, IEEE, 2003, pp. 63–74. `doi:10.1109/CMV.2003.1215004`.

[32] Amazon Mechanical Turk, `https://requester.mturk.com/`, Last accessed: (2020).

[33] C. Reas, B. Fry, Processing: a learning environment for creating interactive web graphics, in: ACM SIGGRAPH 2003 Web Graphics, 2003, pp. 1–1.

[34] M. Lima, The book of trees: Visualizing branches of knowledge, Vol. 5, Princeton Architectural Press New York, 2014.

[35] M. Lind, J. Johansson, M. Cooper, Many-to-many relational parallel coordinates displays, in: 2009 13th International Conference Information Visualisation, IEEE, 2009, pp. 25–31. `doi:10.1109/IV.2009.43`.

[36] J. H. Claessen, J. J. Van Wijk, Flexible linked axes for multivariate data visualization, IEEE Transactions on Visualization and Computer Graphics 17 (12) (2011) 2310–2316. `doi:10.1109/TVCG.2011.201`.

[37] R. Rosenbaum, J. Zhi, B. Hamann, Progressive parallel coordinates, in: 2012 IEEE Pacific Visualization Symposium, IEEE, 2012, pp. 25–32.

[38] G. Palmas, M. Bachynskyi, A. Oulasvirta, H. P. Seidel, T. Weinkauf, An edge-bundling layout for interactive parallel coordinates, in: 2014

IEEE Pacific Visualization Symposium, IEEE, 2014, pp. 57–64. `doi:`
`10.1109/PacificVis.2014.40`.

[39] R. Kanjanabose, A. Abdul-Rahman, M. Chen, A multi-task comparative study on scatter plots and parallel coordinates plots, in: Computer Graphics Forum, Vol. 34, 3, Wiley Online Library, 2015, pp. 261–270. `doi:10.1111/cgf.12638`.

[40] B. C. Kwon, 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, ACM, 2016, pp. 993–997. `doi:10.1145/2858036.2858101`.

[41] High-D Tool, https://www.high-d.com/, Last accessed: (2021).

[42] Mondrian Tool, `http://www.theusrus.de/Mondrian/`, Last accessed: (2021).

[43] Quadrigram Tool, `https://www.quadrigram.com/`, Last accessed: (2021).

[44] Xdat Tool: A free parallel coordinates software, `https://www.xdat.org/`, Last accessed: (2021).

[45] Qualtrics, `https://www.qualtrics.com`, Last accessed: (2020).

[46] M. O. Ward, G. Grinstein, D. Keim, Interactive data visualization: foundations, techniques, and applications, CRC Press, 2010.

[47] GGobi Data Visualization System, http://ggobi.org/, Last accessed: (2021).

[48] Sliver Tool: Multivariate Data Visualization Software, `http://www.sliversoftware.com/features.htm`, Last accessed: (2021).

[49] Microsoft PowerBI, `https://powerbi.microsoft.com/en-us/downloads/`, Last accessed: (2021).

[50] Spotfire Tool, `https://www.tibco.com/products/tibco-spotfire`, Last accessed: (2021).

[51] IBM SPSS Statistics, `https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-26`, Last accessed: (2021).

[52] Xmdv Datasets, `http://davis.wpi.edu/xmdv/datasets.shtml`, Last accessed: (2020).

[53] High-D Datasets, `https://www.high-d.com/datasets/`, Last accessed: (2021).

[54] M. A. Yalçın, N. Elmqvist, B. B. Bederson, Keshif: Rapid and expressive tabular data exploration for novices, IEEE transactions on visualization and computer graphics 24 (8) (2017) 2339–2352.

[55] A. Kirk, Visualizing Data, `https://www.visualisingdata.com/resources/`, Last accessed: (2020).

[56] X. Liu, M. Alharbi, J. Best, J. Chen, A. Diehl, E. Firat, D. Rees, Q. Wang, R. S. Laramee, Visualization resources: A starting point, in: 2021 25th International Conference Information Visualisation (IV), IEEE, 2021, pp. 160–169.

[57] R. Borgo, L. Micallef, B. Bach, F. McGee, B. Lee, Information visualization evaluation using crowdsourcing, in: Computer Graphics Forum, Vol. 37, 3, Wiley Online Library, 2018, pp. 573–595. `doi:10.1111/cgf.13444`.

[58] V. Braun, V. Clarke, Using thematic analysis in psychology, Qualitative research in psychology 3 (2) (2006) 77–101.

Appendix A. Appendix

Post-Experiment Feedback Questionnaire*.*

1. How difficult did you find the questions after the tutorial (post-tutorial)? 1-Not at all, 7-Very much
2. Please explain why you felt the post-tutorial questions are difficult or easy. (Space is given to write)
3. How confident were you answering the post-tutorial questions? 1-Not at all, 7-Very much
4. Did you struggle to answer the post-tutorial questions? 1-Not at all, 7-Very much
5. What would you suggest improving the post-tutorial questions? (Space is given to write)
6. How effective did you find the video tutorial for learning parallel coordinates? 1-Not at all, 7-Very much
7. Do you think you perform better at answering the post-tutorial questions after watching the video tutorial? 1-Not at all, 7-Very much

    The questions below are applicable to the participants who watch the software video tutorial.
8. How helpful was the software tool demonstration? 1-Not at all, 7-Very much
9. Why (or why not) do you think the Parallel Coordinates the software tool demonstration was helpful? (Space is given to write)
10. Was the Parallel Coordinates software tool effective enough to understand the parallel coordinates and the data presented? (Space is given to write)
11. Why (or why not) do you think the Parallel Coordinates software tool was effective? (Space is given to write)

Examples of the Easiest and the Most Difficult PCP Literacy Test Questions*.*
   Q24. This parallel coordinate plot displays information from the 2004 US presidential election that includes records of voters' demographics including the names of three political candidates. The data variables are presented as state name. candidates Bush, Kerry, and Nader the total number of votes,
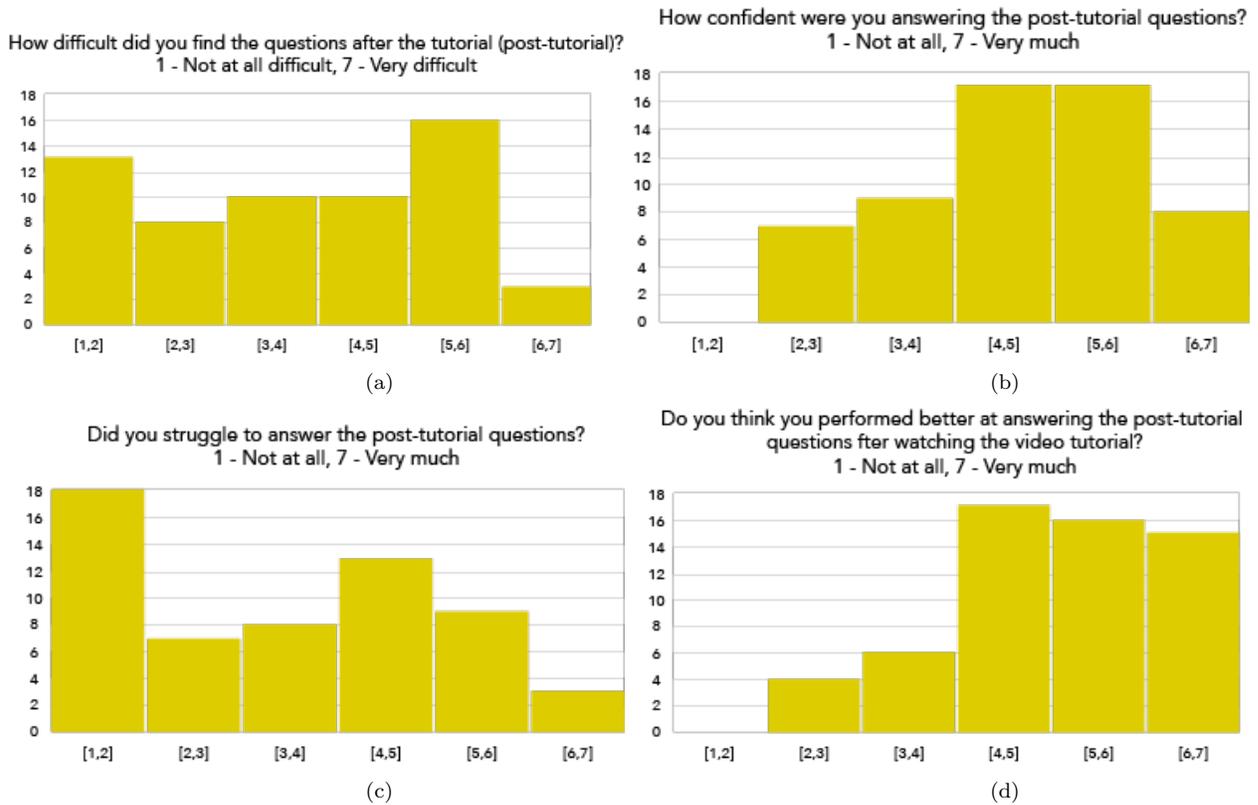
How difficult did you find the questions after the tutorial (post-tutorial)?
1 - Not at all difficult, 7 - Very difficult

(a)

How confident were you answering the post-tutorial questions?
1 - Not at all, 7 - Very much

(b)

Did you struggle to answer the post-tutorial questions?
1 - Not at all, 7 - Very much

(c)

Do you think you performed better at answering the post-tutorial questions fter watching the video tutorial?
1 - Not at all, 7 - Very much

(d)

Figure A.12: Participants' answers to questionnaire

rent in US Dollars, unemployment, and obesity rates. *Which variable has an indirect correlation with the unemployment rate?* (see Figure A.13)

Q26. This image provides information about the US Population from 1950 to 2006 in the thousands. The US Population data variables are plotted on axes such as year, total population, the percent change in the population, resident population, and the civilian population. *The percentage of change started to increase after a certain year.* (see Figure A.14)

Q28. The image plots the characteristics of time-oriented attributes of the planet Venus. The attributes of Venus are presented as date, hour, latitude, longitude, plasma velocity, plasma density, and plasma temperature. *What is the approximate maximum value for velocity?* (see Figure A.15)

Q67. The image shows information about coal-mining disasters for over a century. Each polyline represents one coal disaster with information such as a month of the accident, year of the accident, day of the year, intervals
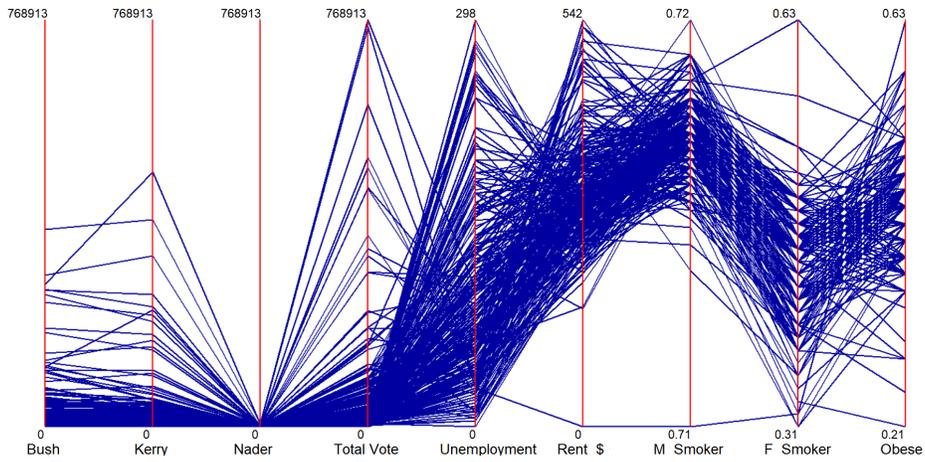
Figure A.13: PCP image of one of the most difficult questions according to ranking of participants' correctly answered questions (see Figure 10 and 11) Options: A) Nadeer B) Rent C) M_Smoker D) Obese E) Not sure F) None of the above

between coal mining disasters, and the number of deaths. *Approximately, what is the range of values for the year attribute?* (see Figure A.16)
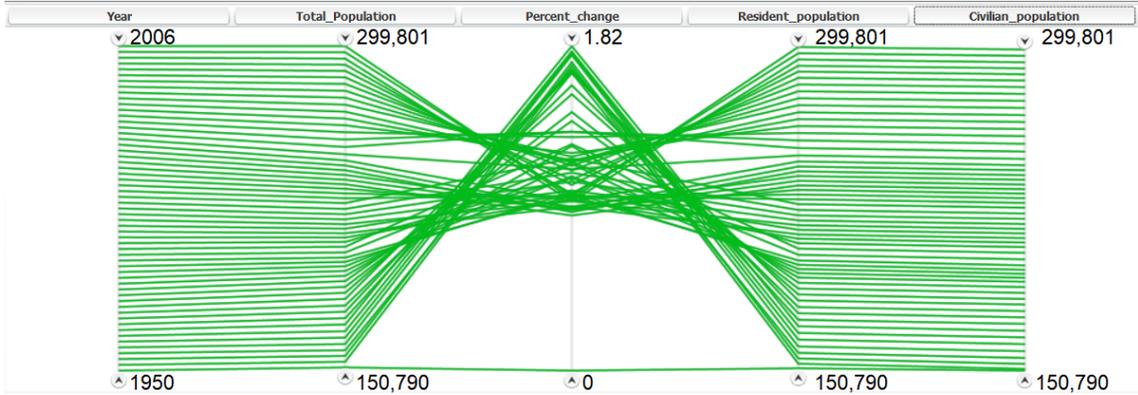
Figure A.14: PCP image of one of the most difficult questions according to ranking of participants' correctly answered questions (see Figure 10 and 11) Options: A) True B) False C) Not sure
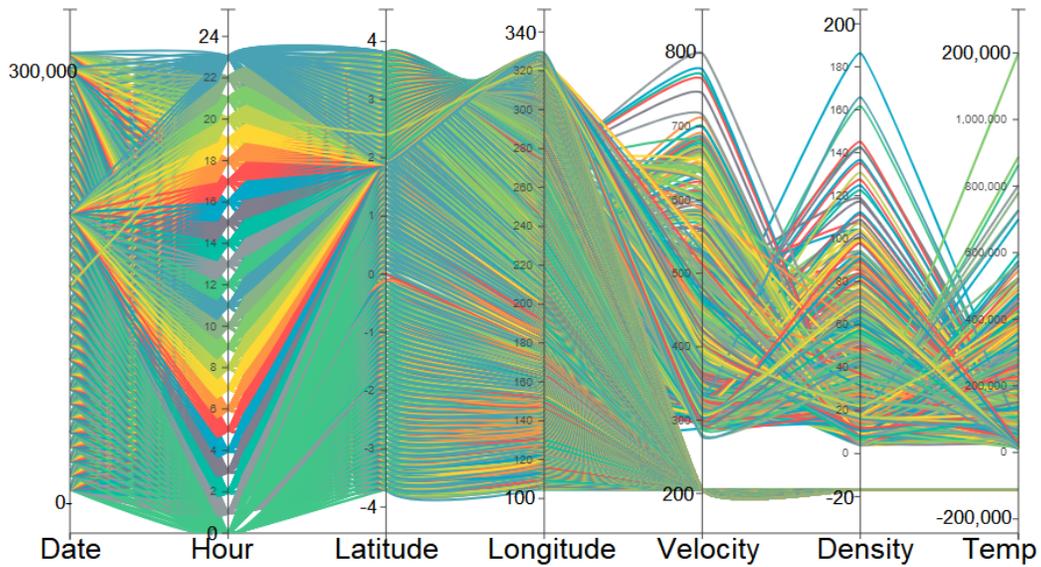

Figure A.15: PCP image of one of the easiest questions according to ranking of participants' correctly answered questions (see Figure 10 and 11) Options: A) 500 B) 600 C)700 D) 800 E) Not sure F) None of the above
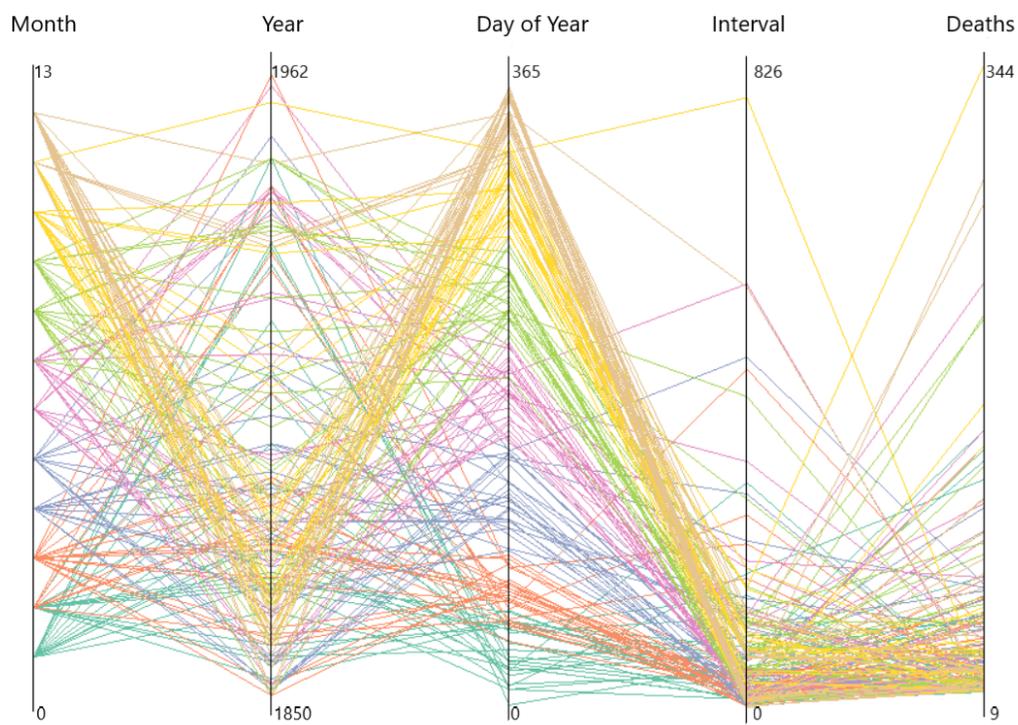
Figure A.16: PCP image of one of the easiest questions according to ranking of partic-
ipants' correctly answered questions (see Figure 10 and 11) Options: A) 1830-1852 B)
1850-1962 C) 1910-1955 D)1910-1995 E) Not sure F) None of the above