A Study on a User-Controlled Radial Tour for Variable Importance in High-Dimensional Data

Nicholas Spyrison, Dianne Cook, Kim Marriott

Abstract—Principal component analysis is a long-standing go-to method for exploring multivariate data. The principal components are linear combinations of the original variables, ordered by descending variance. The first few components typically provide a good visual summary of the data. *Tours* also make linear projections of the original variables but offer many different views, like examining the data from different directions. The grand tour shows a smooth sequence of projections as an animation following interpolations between random target bases. The manual radial tour rotates the selected variable's contribution into and out of a projection. This allows the importance of the variable to structure in the projection to be assessed. This work describes a mixed-design user study evaluating the radial tour's efficacy compared with principal component analysis and the grand tour. A supervised classification task is assigned to participants who evaluate variable attribution of the separation between two classes. Their accuracy in assigning the variable importance is measured across various factors. Data were collected from 108 crowdsourced participants, who performed two trials with each visual for 648 trials in total. Mixed model regression finds strong evidence that the radial tour results in a large increase in accuracy over the alternatives. Participants also reported a preference for the radial tour in comparison to the other two methods.

Index Terms—Multivariate data visualization, variable importance, radial tour, linear dimension reduction,

✦

1 INTRODUCTION

Despite decades of research, multivariate data continues to provide fascinating challenges for visualization. Data visualization is important because it is a key element of exploratory data analysis [43] for assessing model assumptions and as a cross-check on numerical summarization [2, 26, 50]. One challenge is measuring if a new technique yields a more informed perception of information than current practices.

Dimension reduction is commonly used with visualization to provide informative low-dimensional summaries of quantitative multivariate data. Principal component analysis (PCA) [34] is one of the first methods ever developed, and it remains very popular. Visualization of PCA is typically in the form of static scatterplots of a few leading components. When the scatterplot is accompanied by a visual representation of the basis they are called a biplot [17]. A basis is a $p \times d$ matrix of the linear combination of the p variables mapped to a smaller d-dimensional space. That is, it is an orthogonal rotation matrix, the magnitude, and the angle that the variables contribute.

Dynamic visualizations called *tours* [4] animate through a sequence of linear projections (orthonormal bases). Instead of a static view, tours provide a smoothly changing view by interpolating between bases. There are various types of tours distinguished by how the paths are generated. Asimov originally animated between randomly selected bases in the *grand* tour. The *manual* tour [11] allows for user control over the basis changes. A selected variable (or component) can be rotated into or out of view or to a particular value. The *radial tour* [42] is a variant of the manual tour that fixes the contribution angle and changes

the magnitude along the radius. The permanence of the data points from basis to basis holds information between intermediate interpolated projections, and the user control of the basis could plausibly lead to more information being perceived than a static display. This is a hypothesis that a user study could assess.

Empirical studies have rarely assessed tours. An exception is [31], who compares scatterplots of grand tours on 2D monitors with 3D (stereoscopic, not head-mounted) over $n = 15$ participants. Participants perform cluster detection, dimensionality estimation, and radial sparseness tasks on six-dimensional data. They find that stereoscopic 3D leads to more accuracy in cluster identification, though the time to interact with the display was much higher in the 3D environment. In this work, we extend the evaluation of tours which compares the radial tour as benchmarked against the grand tour and discrete pairs of principal components.

The contribution of this paper is an empirical user study comparing the radial tour against PCA and the grand tour for assessing variable attribution on clustered data. This is the first empirical evaluation of the radial or manual tour. We discuss how this fits with other multivariate data visualization techniques and coordinated views of linear projections.

We are particularly interested in assessing the effectiveness of the new radial tour relative to common practice with PCA and grand tour. The user influence over a basis, uniquely available in the radial tour, is crucial to testing variable sensitivity to the structure visible in projection. If the contribution of a variable is reduced and the feature disappears, then we say that the variable was sensitive to that structure. For example, Fig. 1 shows two projections of simulated data. Panel (a) has identified the separation between the two clusters. The contributions in panel (b) show no such cluster separation. The former has a large contribution of V2 in the direction of separation, while it is negligible in the right frame. Because of this, we say that V2 is sensitive to the separation of the clusters.

Variable sensitivity is important for the interpretation of machine learning models. They are the magnitude and direction of contribution to the model. It is important that developers maintain the interpretability of models. Exploratory Artificial Intelligence (XAI) [1, 3], is an emerging field that extends the interpretability of such black-box models. Multivariate data visualization is essential for exploring feature spaces and communicating interpretations of models [5, 6, 47].

The paper is structured as follows. Sect. 2 provides background on standard visualization methods and linear dimension reduction techniques. Sect. 3 describes the experimental factors, task, and accuracy measure used. The results of the study are discussed in Sect. 4. Con-

- *Monash University*
  Australia
  nicholas.spyrison@monash.edu
  ORCiD: 0000-0002-8417-0212.
- *Monash University*
  Australia
  dicook@monash.edu
  ORCiD: 0000-0002-3813-7155
- *Monash University*
  Australia
  kim.marriott@monash.edu
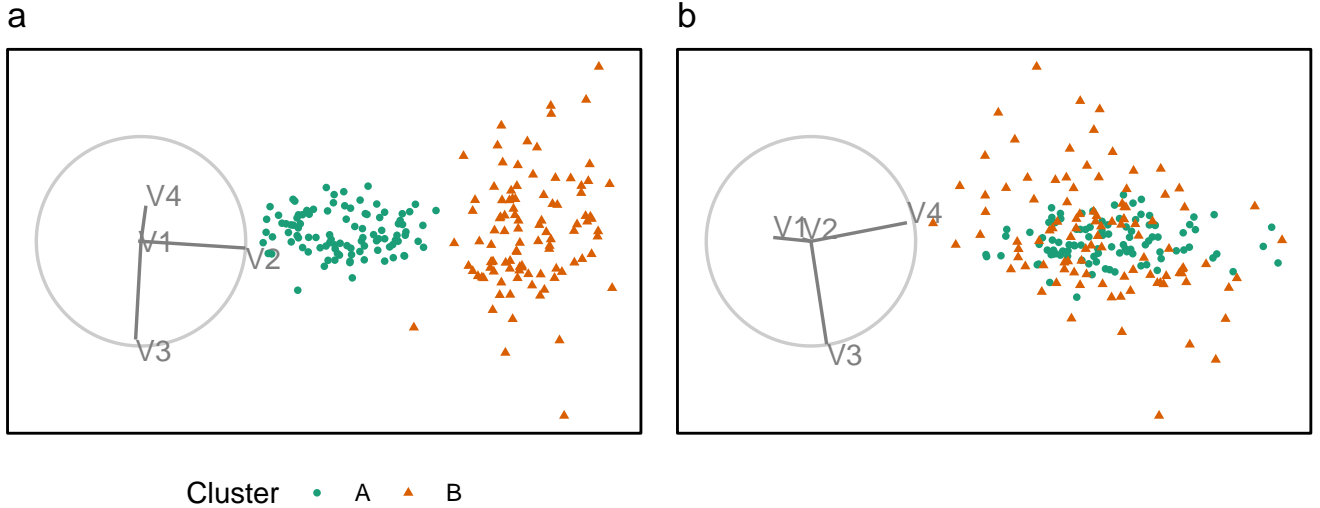 *ORCiD: 0000-0002-9813-0377*

Fig. 1. Illustration of cluster separation affected by variable importance. Panel (a) is a projection mostly of V2 and V3, and the separation between clusters is in the direction of V2, not V3. This suggests V2 is important for clustering, but V3 is not. Panel (b) shows a projection of mostly V3 and V4, with no contribution from V2 and little from V3. That there is no separation between the clusters indicates that V3 and V4 are not important.

clusions and potential future directions are discussed in Sect. 6. More results, participant demographics, and analysis of the response time are available in the Supplemental Materials.

2 RELATED WORK

Consider the data to be a matrix of n observations (rows) and p variables (columns), denoted as $X_{n \times p}$.

2.1 Orthogonal multivariate visualization

Grinstein [19] illustrates many multivariate visualization methods. In particular, this work shows examples of actual visuals. Liu [25] give a good classification and taxonomy of such methods. The content below focuses on the most common visuals that use the full data space before discussing linear combinations of those variables in projections.

2.1.1 Scatterplot matrix

One could consider looking at p histograms or univariate densities. Doing so will miss features in two or more dimensions. Fig. 2 shows a scatterplot matrix [9] of the four principal components of simulated data. Such displays do not scale well with dimensions because each plot would get less and less space. Scatterplot matrices can only display information in two orthogonal dimensions, so features in three dimensions may not be fully resolved.

2.1.2 Parallel coordinates plot

Another common way to display multivariate data is with a parallel coordinates plot [32]. Parallel coordinates plots scale well with dimensions but poorly with observations as the lines overcrowd the display. Parallel coordinate plots are asymmetric across variable ordering. In that, shuffling the order of the variable can lead to different conclusions. Another shortcoming is the graphical channel used to convey information. [29] suggests that position is the visual channel that is most perceptible to humans. In the case of parallel coordinates plots, the horizontal axes span variables rather than the values of one variable, causing the loss of a display dimension to be used by our most perceptible visual channel.

2.2 Multivariate projections

At some point, visualization will be forced to turn to dimension reduction to scale better with the dimensionality of the data. Below we

introduce linear projections and the common principal component analysis. Then we touch on nonlinear projections and exclude them from consideration.

2.2.1 Linear

Let data, X, contain n observations of p variables. A linear projection maps a higher p-dimensional space onto a smaller d-space with an affine mapping (where parallel lines stay parallel). A projection, Y, is the resulting space of the data multiplied by a *basis*, A, such that $Y_{n \times d} = X_{n \times p} \times A_{p \times d}$. This is essentially a reorientation of the original variable. This intuition is conveyed by thinking of a shadow as a 2D projection of a 3D object. Rotating the object changes the shadow it casts and, correspondingly, the basis that maps the reorientation of the object.

2.2.2 Principal component analysis

PCA is a good baseline of comparison for linear projections because of its frequent and broad use across disciplines. PCA [34] defines new components, linear combinations of the original variables, ordered by decreasing variation through the help of eigenvalue matrix decomposition. While the resulting dimensionality is the same size, the benefit comes from the ordered nature of the components. The data can be said to be approximated by the first several components. The exact number is subjectively selected given the variance contained in each component, typically guided from a scree plot [8]. Features with sizable signal regularly appear in the leading components that commonly approximate data. However, this is not always the case and component spaces should be fully explored to look for signal in components with less variation. This is especially true for cluster structure [14].

2.2.3 Nonlinear

Nonlinear transformations bend and distort spaces that are not entirely accurate or faithful to the original variable space. Popular modern methods include t-SNE and UMAP [28, 44]. There are various quality metrics, such as Trustworthiness, Continuity, Normalized stress, and Average local error, have been introduced to describe the distortion of the space [16, 18]. Unfortunately, these distortions are hard to visualize and comprehend, effectively breaking the variable interpretability of the resulting space. The intuition of this can be demonstrated with map projections. Snyder [41] lists over 200 different projections that
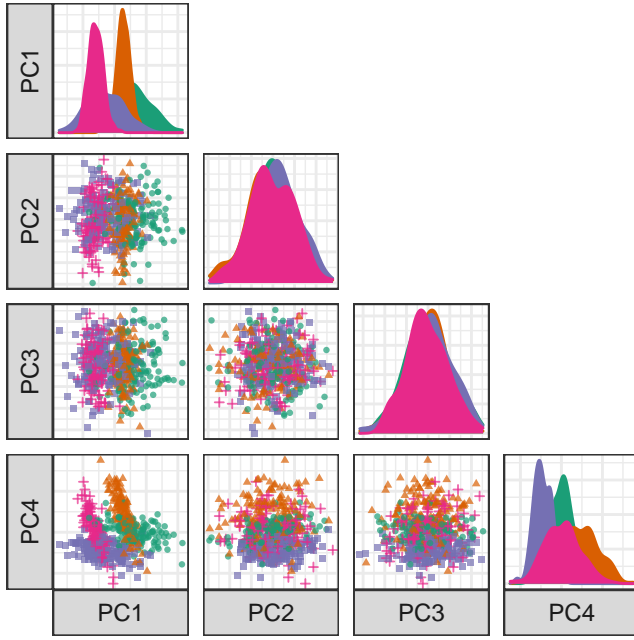
Fig. 2. Scatterplot matrix of the first four principal components of 6D simulated data containing four classes. The separation between classes is primarily in PC1 and PC4. This is not uncommon because PCA is summarizing variance, not cluster structure.

distort the surface of the earth to display as a 2D map, each with unique properties and use cases.

Because of the difficulty of interpreting the distortions of nonlinear spaces and the added subjectivity of hyperparameter selection, we exclude nonlinear techniques and instead, decide to compare three linear techniques.

2.3 Tours, animated linear projections

A *tour* animates through many linear projections. One of the insightful features of the tour is the permanence of the data points; one can track the relative changes of observations as the basis changes, as opposed to discretely jumping to an orthogonal view angle with no intermediate information. Types of tours are distinguished by the generation of their basis paths [13, 22]. In contrast with the discrete orientations of PCA, we compare continuous linear projection changes with grand and radial tours.

2.3.1 Grand tours

Target bases are selected randomly in a grand tour [4]. These target bases are then geodesically interpolated for a smooth, continuous path. The grand tour is the first and most widely known tour. The random selection of target bases makes it a general unguided exploratory tool. The grand tour will make a good comparison that has a continuity of data points similar to the radial tour but lacks the user control enjoyed by PCA and radial tours.

2.3.2 Manual and radial tours

Whether an analyst uses PCA or the grand tour, cannot influence the basis. They cannot explore the structure identified or change the contribution of the variables. User-controlled steering is a key aspect of *manual* tours that helps to test variable attribution.

The manual tour [11] defines its basis path by manipulating the basis contribution of a selected variable. A manipulation dimension is appended onto the projection plane, giving a full contribution to the selected variable. The target bases are then chosen to rotate this newly created manipulation space. This manipulation space is similarly orthogonally restrained. The data is projected through its interpolated

basis and rendered into an animation. When the contribution of one variable changes, the contributions of the other variables must also change, to maintain the orthonormality of the basis. A key feature of the manual tour is that it allows users to control the variable contributions to the basis. Such manipulations can be queued in advance or selected in real time for human-in-the-loop analysis [21]. Manual navigation is relatively time-consuming due to the vast volume of resulting view space and the abstract method of steering the projection basis. First, it is advisable to identify a basis of particular interest and then use the manual tour as a more directed, local exploration tool to explore the sensitivity of a variable's contribution to the feature of interest.

To simplify the task and keep its duration realistic, we consider a variant of the manual tour called a *radial* tour. In a radial tour, the magnitude of along the radius with a fixed angle of contribution, as seen in Fig. 3. The radial tour benefits from both continuity of the data alongside grand tours and user-steering via choosing the variable to rotate.

Manual tours have been recently made available in the **R** package **spinifex** [42], which facilitates manual tour (and radial variant). It also provides an interface for a layered composition of tours and exporting to gif and mp4 with **gganimate** [35] or html widget with **plotly** [40]. It is also compatible with tours made by **tourr** [48]. Now that we have a readily available means to produce various tours, we want to see how they fare against traditional discrete displays commonly used with PCA.

2.4 Other animated linear projections

The work of [15] allows users to interactively change the face of a local display by navigating to adjacent faces on a global overview scatterplot matrix. This offers analysts a way to geometrically explore the transition between adjacent faces of a scatterplot matrix as though rotating the face of dice at right angles. The interpolated bases between the orthogonal faces display linear combinations of three variables at varying degrees. This is what [27] called a *little tour* with the addition of user control. It is a particular type of manual tour where only horizontal or vertical rotation is allowed.

Star Coordinates [20] also arrive at the biplot scatterplot displays starting from the perspective of radial parallel coordinates. [23] extend this idea, mapping it back to orthogonal projections. They provide a means to interpolate through PCA components, the orthogonal contributions of the scatterplot matrix, and the grand tour. This work also defines user-controlled interaction, similar to small steps in a manual or radial tour.

TripAdvisor [30] is an interactive application that plans sequential interpolation between distant target bases. It also provides an additional global context of a subset of possible frames with glyph representation and an overview of variable attribution by summarizing the top ten principal components. It allows for user-steering by using a "touchpad polygon". This touchpad allows for contribution magnitudes to be changed. This is similar to an incremental change with the manual tour.

The number of orthogonal axes in static plots as well as the number of bases to view in a tour increase quadratically with the dimensions, p. This is why it is particularly important to properly select variables or otherwise reduce the dimensions before viewing. PCA, Linear discriminant analysis and entropy are common approaches to variable selection [37, 38, 46]. Such methods often yield a sort of screeplot [8] where the analyst selects a subjective, but informed, number of components to approximate the data while discarding the least information. The variable sensitivity we test for, in contrast, is the act of visual analysis of one variable's contribution to the structure. In practice, this is a tool for the analyst to fine-tune their variable selection or otherwise evaluate the resulting approximated space.

In order to further mitigate the view time, objective functions can be used to inform static or animated biplots. A dissimilarity statistic can be used to solve a basis path for showing a particularly interesting tour [24]. More generally projection pursuit can be used to conduct a guided tour of any objective function applied to an embedding space [12, 13]. However, the function optimized is likely to show some feature of interest if it is ultimately selected by the analyst. The ability to stop
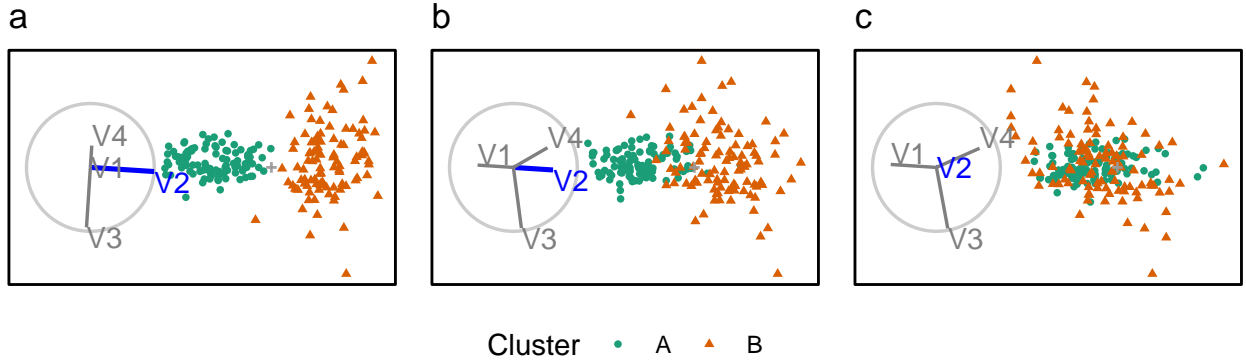
a   b   c

Cluster   •   A   ▲   B

Fig. 3. A radial tour changing the contribution of $V2$. The contribution is in the direction of cluster separation. When its contribution is removed, the clusters overlap (right). Because of this, we say that $V2$ is sensitive to the separation of these two species.

and control the exploration at any point only stands to improve one's understanding of the data.

2.5 Empirical evaluation

Some studies compare visualizations across complete contributions of variables. Chang [10] conducted an $n = 51$ participant study comparing parallel coordinate plots and scatterplot matrix either in isolation, sequentially, or as a coordinated view. Accuracy, completion time, and eye focus were measured for six tasks. Three tasks were more accurate with scatterplot matrix and three with parallel coordinates, while the coordinated view was usually marginally more accurate than the max of the separate visuals. Cao [7] compare nonstandardized line-glyph and star-glyphs with standardized variants (with and without fill under the curve). Each of the $n = 18$ participants performed 72 trials across the six visuals, two levels of dimensions, and two levels of observations. Visuals with variable standardization outperformed the nonstandardized variants, and the radial star-glyph reportedly outperformed the line variant.

Other studies have investigated the relative benefits of projecting to 2- or 3D scatterplots in PCA-reduced spaces. Gracia [18] conducted an $n = 40$ user study comparing 2- and 3D scatterplots on traditional 2D monitors. Participants perform point classification, distance perception, and outlier identification tasks. The results are mixed and primarily have small differences. There is some evidence to suggest a lower error in distance perception from a 3D scatterplot. Wagner Filho [45] performed an $n = 30$ mixed-design study on PCA reduced space using scatterplot displays between 2D on monitors, 3D on monitors, and 3D display with a head-mounted display. None of the tasks on any dataset lead to a significant difference in accuracy. However, the immersive display reduced effort and navigation, resulting in higher perceived accuracy and engagement. Sedlmair [39] instead used two expert coders to evaluate 75 datasets and four dimension reduction techniques across the displays of 2D scatterplots, interactive 3D scatterplots, and 2D scatterplot matrices. They suggested a tiered guidance approach finding that 2D scatterplots are often sufficient to resolve a feature. If not, try 2D scatterplots on a different dimension reduction technique before going to scatterplot matrix display or concluding a true negative. They find that interactive 3D scatterplots help in very few cases.

2.6 Conclusion

Orthogonal axes visualizations either scale poorly with dimensionality or introduce an asymmetry of the variable ordering. Projections visualize the full p-data as fewer dimensions, traditionally 1-3 at a time. In linear, orthogonal projections, the resulting space is composed of a linear combination of the original variables that maintain variable interpretability. While nonlinear techniques distort and bend space in different ways that are hard to visualize and communicate.

Tours are linear projections that are animated over changes in the basis. Several more-recent, orthographic-star coordinate methods independently reach animated linear projections similar to tours. Some quality metrics and empirical studies compare techniques but scarcely with animated methods. Below we conduct a user study to compare the radial tour with PCA and the grand tour on a variable attribution task on clustered data.

3 USER STUDY

The experiment was designed to assess the performance of the radial tour relative to the grand tour and PCA for interpreting the variable attribution to the separation between two clusters. Data were simulated across three experimental factors: location of the cluster separation, cluster shape, and data dimensionality. Participant responses were collected using a web application and crowdsourced through prolific.co, [33] an alternative to MTurk.

3.1 Objective

PCA will be used as a baseline for comparison as it is the most commonly used linear embedding. It will use static, discrete jumps between orthogonal components. The grand tour will act as a secondary control that will help evaluate the benefit of observation trackability between nearby animation bases but without user-control of its path. Lastly, the radial tour will be compared, which benefits from the continuity of animation and user control of the basis.

Then for some subset of tasks, we expect to find that the radial tour performs most accurately. Conversely, there is less to be certain about the accuracy of such limited grand tours as there is no objective function in selecting the bases; it is possible that the random selection of the target bases altogether avoids the bases showing cluster separation. However, given that the data dimensionality is modest, it is probable that the grand tour coincidentally regularly crossed bases with the correct information for the task.

Experimental factors and the definition of an accuracy measure are given below. The null hypothesis can be stated as:

H_0 : accuracy does not change across the visual methods

H_α : accuracy does change across the visual methods

3.2 Visual factors

The visual methods are tested mixed design, with each visual being evaluated twice by each participant. Scatterplot matrices or parallel coordinates could alternatively be used to visualize these spaces. However, we opt to focus on single biplot displays to focus on the differences between the radial tour and its most comparable visuals, rather than a comprehensive comparison of visual methods. The rest of this section
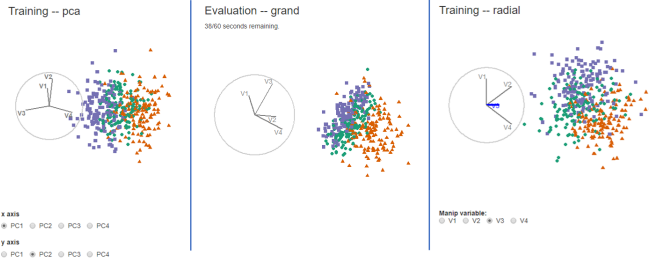
Fig. 4. Examples of the application displays for PCA, grand tour, and radial tour.

discusses the design standardization and unique input associated with each visual.

The visualization methods were standardized wherever possible. Data were displayed as 2D scatterplots with biplots. All aesthetic values (color-blind safe colors, shapes, sizes, absence of legend, and axis titles) were constant. The variable contribution biplot was always shown left of the scatterplot embeddings with their aesthetic values consistent. What did vary between visuals were their inputs.

PCA allowed users to select between the top four principal components for each axis regardless of the data dimensionality (four or six). Upon changing an axis, the visual would change to the new view of orthogonal components without displaying intermediate bases. There was no user input for the grand tour; users were instead shown a 15-second animation of the same randomly selected path (variables containing cluster separation were shuffled after simulation). Participants could view the same clip up to four times within the time limit. Radial tours allowed participants to select the manipulation variable. The starting basis was initialized to a half-clock design, where the variables were evenly distributed in half of the circle. This design was created to be variable agnostic while maximizing the independence of the variables. Selecting a new variable resets the animation where the new variable is manipulated to a complete contribution, zeroed contribution, and then back to its initial contribution. Animation and interpolation parameters were held constant across grand and radial tour (five bases per second with a step size of 0.1 radians between interpolated bases). Fig. 4 displays screen captures of the visuals in the application.

3.3 Experimental factors

In addition to the visual method, data are simulated across three experimental factors. First, the *location* of the separation between clusters is controlled by mixing a signal and a noise variable at different ratios. Secondly, the *shape* of the clusters reflects varying distributions of the data. And third, the *dimension*-ality of the data is also tested. The levels within each factor are described below, and Fig. 5 gives a visual representation.

The *location* of the separation between the clusters is at the heart of the measure. It would be good to test a few varying levels. To test the sensitivity, a noise and signal variable are mixed at different ratios. The separation between clusters is mixed at the following percentages: 0/100% (not mixed), 33/66%, 50/50% (evenly mixed).

In selecting the *shape* of the clusters, the convention given by Scrucca et al. (2016) is followed. They describe 14 variants of model families containing three clusters. The model family name is the abbreviation of the clusters' respective volume, shape, and orientation. The levels are either *E*qual or *V*ary. The models EEE, EEV, and EVV are used. For instance, in the EEV model, the volume and shape of clusters are constant, while the shape's orientation varies. The EVV model is modified by moving four-fifths of the data out in a ">" or banana-like shape.

Dimension-ality is tested at two modest levels: four dimensions containing three clusters and six with four clusters. Such modest dimensionality is required to limit the difficulty and search space to make the task realistic for crowdsourcing.
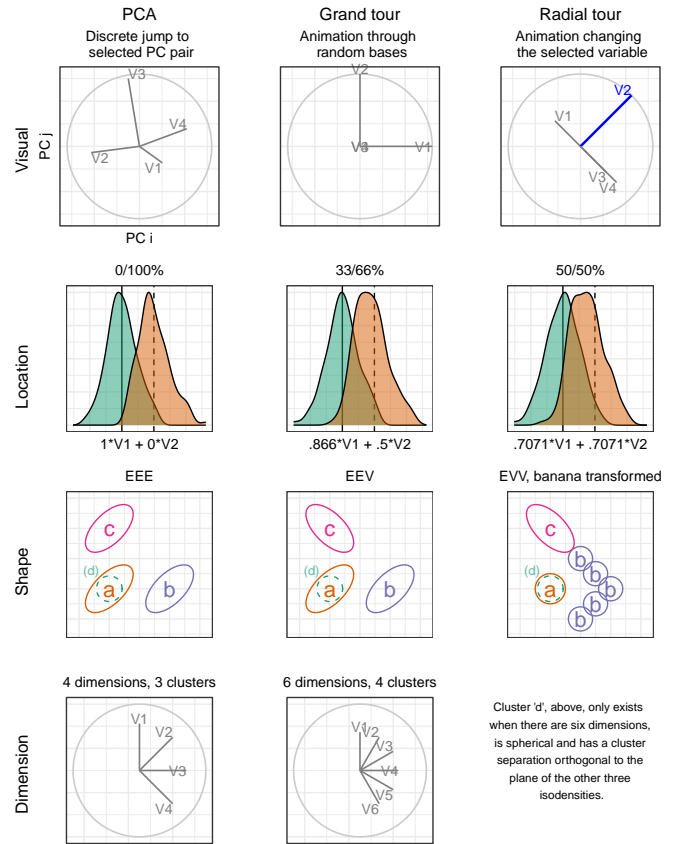

Fig. 5. Levels of the visuals and three experimental factors: location of cluster separation, the shape of clusters, and dimensionality of the sampled data.

5

3.4 Task and evaluation

With our hypothesis formulated, let us turn our attention to the task and how it is evaluated. Participants were asked to "check any/all variables that contribute more than average to the cluster separation of the green circles and the orange triangles". This was further explained in the explanatory video as "mark any and all variable that carries more than their fair share of the weight, or one quarter in the case of four variables". The participant instruction video can be viewed at https://vimeo.com/712674984.

The instructions iterated several times in the video were: 1) use the input controls to find a basis that contains separation between the clusters of green circles and orange triangles, 2) look at the orientation of the variable contributions in the grey circle (biplot axes orientation), and 3) select all variables that contribute more than uniformed distributed cluster separation in the scatterplot. Independent of the experimental level, participants were limited to 60 seconds for each evaluation of this task. This restriction did not impact many participants as the 25th, 50th, and 75th quantiles of the response time were about 7, 21, and 30 seconds, respectively.

The accuracy measure of this task was designed with a couple of features in mind. 1) Symmetric about the expected value, without preference for under- or over-guessing. 2) Heavier than linear weight with an increasing difference from the expected value. The following measure is defined for evaluating the task.

Let the data $\mathbf{X}_{ijk}, i = 1, ..., n; j = 1, ..., p; k = 1, ..., K$ be simulated observations containing clusters of observations of different distributions. Where n is the number of observations, p is the number of variables, and K indicates the number of clusters. Cluster membership is exclusive; an observation cannot belong to more than one cluster.

The weights, w, is a vector of the variable-wise difference between the mean of two clusters of less $1/p$, the expected cluster separation if it were uniformly distributed. Accuracy, A is defined as the signed square of these weights if selected by the participant. Participant responses are a logical value for each variable — whether or not the participant thinks each variable separates the two clusters more than uniformly distributed separation. Weights comparing clusters 1 and 2 are calculated as follows:

$$A = \sum_{j=1}^{p} I(j) \cdot sign(w_j) \cdot w_j^2, \text{ where}$$

$$w_j = \frac{\overline{X}_{\cdot j1} - \overline{X}_{\cdot j2}}{\sum_{j=1}^{p}(|\overline{X}_{\cdot j1} - \overline{X}_{\cdot j2}|)} - \frac{1}{p}$$

where
I is the indicator function, returning a binary response
$\overline{X}_{\cdot jk}$, mean of the j-th variable of the k-th cluster

where $I(j)$ is the indicator function, the binary response for variable j. Fig. 6 shows one projection of a simulation with its observed variable separation (wide bars), expected uniform separation (dashed line), and accuracy if selected (thin vertical lines).

3.5 Randomized factor assignment

Now, with simulation and their artifacts in hand, this section covers how the experimental factors are assigned and demonstrate how this is experienced from the participant's perspective.

The study is sectioned into three periods. Each period is linked to a randomized level of visual and location. The order of dimension and shape are of secondary interest and are held constant in increasing order of difficulty; four then six dimensions and EEE, EEV, then EVV-banana, respectively.

Each period starts with an untimed training task at the simplest remaining experimental levels; location = 0/100%, shape = EEE, and four dimensions with three clusters. This serves to introduce and familiarize participants with input and visual differences. After the training, the participant performs two trials with the same visual and

location level across the increasing difficulty of dimension and shape. The plot was removed after 60 seconds, though participants rarely reached this limit.

We assigned these factors based on the following order: visual methods, location, shape, and dimensionality. We first assigned three visual methods to three different sessions. The session order and the order of location follow a nested Latin square. The order of dimension and shape are assigned based on increasing order of difficulty.

Through pilot studies sampled by convenience (information technology and statistics Ph.D. students attending Monash University), it was estimated that three complete evaluations are needed to power the study properly, a total of $N = 3 \times 3!^2 = 108$ participants.

3.6 Participants

$N = 108$ participants were recruited via prolific.co (Palan and Schitter 2018). Participants are restricted based on their claimed education requiring that they have completed at least an undergraduate degree (some 58,700 of the 150,400 users at the time). This restriction is used on the premise that linear projections and biplot displays will not be regularly used for consumption by general audiences. There is also the implicit filter that Prolific participants must be at least 18 years of age and have implicit biases of timezone, internet availability, language compatibility, and socioeconomic status. Participants were compensated for their time at £7.50 per hour, whereas the mean duration of the survey was about 16 minutes. Previous knowledge or familiarity was minimal, as validated in the follow-up survey. The Supplemental Materials include a heatmap distribution of age and education paneled across preferred pronouns of the participants that completed the survey, who are relatively young, well-educated, and slightly more likely to identify as males.

4 RESULTS

To recap, the primary response variable is accuracy, as defined in Sect. 3.4. Two primary data sets were collected; the user study evaluations and the post-study survey. The former is the 108 participants with the experimental factors: visual, location of the cluster separation signal, the shape of the variance-covariance matrix, and the dimensionality of the data. Experimental factors and randomization were discussed in Sect. 3.3. A follow-up survey was completed by 84 of these 108 people. It collected demographic information (preferred pronoun, age, and education) and subjective measures for each visual (preference, familiarity, ease of use, and confidence).

Below a battery of mixed regression models is built to examine the degree of the evidence and the size of the effects of the experimental factors. Then, Likert plots and rank-sum tests to compare the subjective measures between the visuals.

4.1 Accuracy

To quantify the contribution of the experimental factors to the accuracy, mixed-effects models were fit. All models have a random effect term on the participant and the simulation. These terms explain the amount of error attributed to the individual participant's effect and variation due to the random sampling data.

In building a set of models to test, a base model with only the visual term being compared with the full linear model term and progressively interacting with an additional experimental factor. The models with three and four interacting variables are rank deficient; there is not enough varying information in the data to explain all interacting terms.

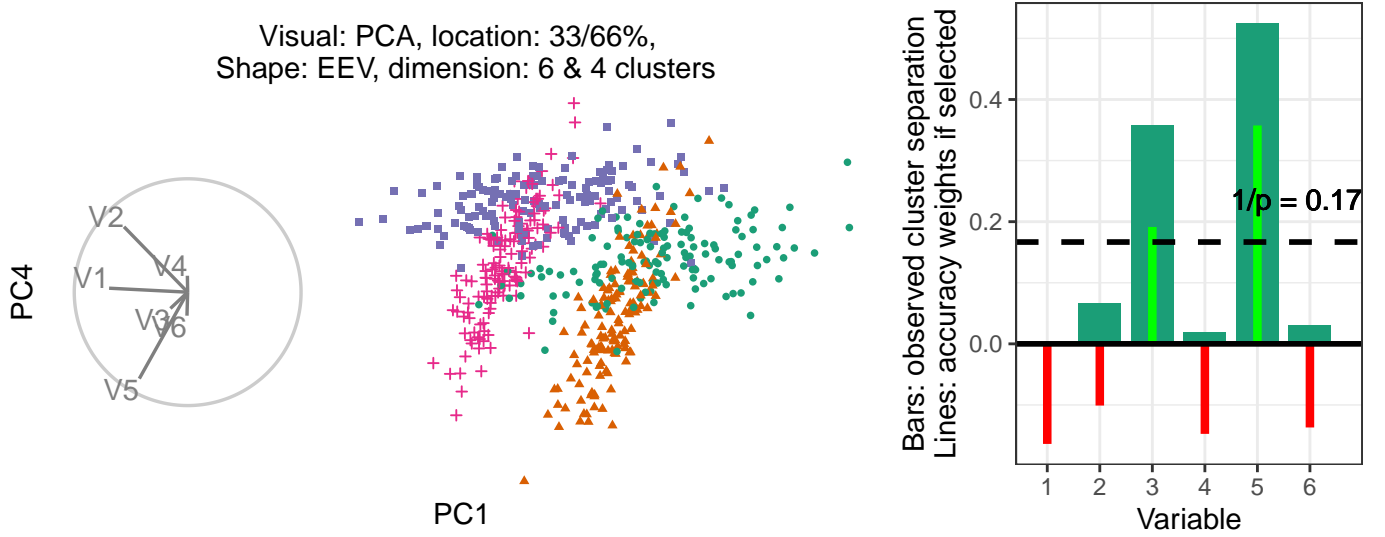| Fixed effects	Full model
α	$\widehat{Y} = \mu + \alpha_i + \mathbf{Z} + \mathbf{W} + \varepsilon$
$\alpha + \beta + \gamma + \delta$	$\widehat{Y} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_l + \mathbf{Z} + \mathbf{W} + \varepsilon$
$\alpha \times \beta + \gamma + \delta$	$\widehat{Y} = \mu + \alpha_i \times \beta_j + \gamma_k + \delta_l + \mathbf{Z} + \mathbf{W} + \varepsilon$
$\alpha \times \beta \times \gamma + \delta$	$\widehat{Y} = \mu + \alpha_i \times \beta_j \times \gamma_k + \delta_l + \mathbf{Z} + \mathbf{W} + \varepsilon$
$\alpha \times \beta \times \gamma \times \delta$	$\widehat{Y} = \mu + \alpha_i \times \beta_j \times \gamma_k \times \delta_l + \mathbf{Z} + \mathbf{W} + \varepsilon$

Fig. 6. Illustration of how accuracy is measured. (L), Scatterplot and biplot of PC1 by PC4 of a simulated data set (R) illustrate cluster separation between the green circles and orange triangles. Bars indicate observed cluster separation, and (red/green) lines show the accuracy of the variable if selected. The horizontal dashed line has a height $1/p$, the expected value of cluster separation. The accuracy weights equal the signed square of the difference between each variable value and the dashed line.

Table 1. Model performance of random effect models regressing accuracy. Complex models perform better in terms of R^2 and RMSE, yet AIC and BIC penalize their large number of fixed effects in favor of the much simpler model containing only the visuals. Conditional R^2 includes error explained by the random effects, while marginal does not.

Model	AIC	BIC	R2 cond.	R2 marg.	RMSE
a	**-71**	**-71**	-44.219	0.303	0.289
a+b+c+d	-45	-45	4.063	0.334	0.294
a*b+c+d	-26	-25	41.445	0.338	0.293
a*b*c+d	28	32	167.092	**0.383**	0.309
a*b*c*d	105	116	**360.052**	0.37	**0.19**

where
> μ is the intercept of the model
> α_i is the visual | $i \in$ (pca, grand, radial)
> β_j is the location | $j \in$ (0/100, 33/66, 50/50% mix)
> γ_k is the shape | $k \in$ (EEE, EEV, EVV banana)
> δ_l is the dimension | $l \in$ (4 & 3, 6 & 4) var & clusters
> $\mathbf{Z} \sim \mathcal{N}(0, \tau)$ is the random effect of participant
> $\mathbf{W} \sim \mathcal{N}(0, \upsilon)$ is the random effect of simulation
> $\varepsilon \sim \mathcal{N}(0, \sigma)$ is the remaining error of the model

Table 1 compares the model summaries across increasing complexity. The $\alpha \times \beta + \gamma + \delta$ model is selected to examine in more detail as it has relatively high condition R^2 and not overly complex interacting terms. Table 2 looks at the coefficients for this model. There is strong evidence suggesting a relatively large increase in accuracy from the radial tour, though there is evidence that almost all of the increase the is lost under 33/66% mixing.

We also want to visually examine the conditional variables in the model. Fig. 7 illustrates the accuracy for each model term shown as mean and 95% confidence interval.

4.2 Subjective measures

Modeling has proven that the use of the radial tour leads to a sizable improvement in the accuracy measure for this task. This is not the whole story. It is desirable to know what the users think of using the visuals. We follow the direction set by [45]. They observe four

Table 2. The task accuracy model coefficients for $\widehat{Y} = \alpha \times \beta + \gamma + \delta$, with visual = pca, location = 0/100%, shape = EEE, and dim = 4 held as baselines. Visual being radial is the fixed term with the strongest evidence supporting the hypothesis. Interacting with the location term, there is evidence suggesting radial performs with minimal improvement for 33/66% location mixing.

|	Est	SE	df	t val	Prob	
(Intercept)	0.10	0.06	16.1	1.54	0.143	
Factor						
VisGrand	0.06	0.04	622.1	1.63	0.104	
VisRadial	0.14	0.04	617.0	3.77	0.000	***
Fixed effects						
Loc33/66%	-0.02	0.07	19.9	-0.29	0.777	
Loc50/50%	-0.04	0.07	20.0	-0.66	0.514	
ShapeEEV	-0.05	0.06	11.8	-0.82	0.427	
ShapeBanana	-0.09	0.06	11.8	-1.54	0.150	
Dim6	-0.01	0.05	11.8	-0.23	0.824	
Interactions						
VisGrand:Loc33/66	-0.02	0.06	588.9	-0.29	0.774	
VisRadial:Loc33/66	-0.12	0.06	586.5	-2.13	0.033	*
VisGrand:Loc50/50	-0.03	0.06	591.6	-0.47	0.641	
VisRadial:Loc50/50	-0.06	0.06	576.3	-1.16	0.248	
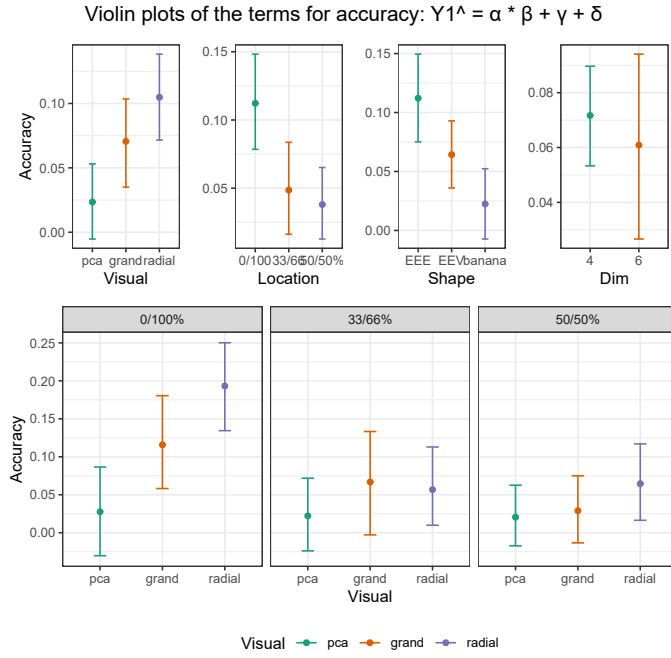
Fig. 7. Accuracy of terms of the model $\widehat{Y} = \alpha \times \beta + \gamma + \delta$. Viewing the marginal accuracy of the terms corroborates the primary findings that the use of the radial tour leads to a significant increase in accuracy, at least over PCA, and this effect is particularly well supported when no location mixing is applied.
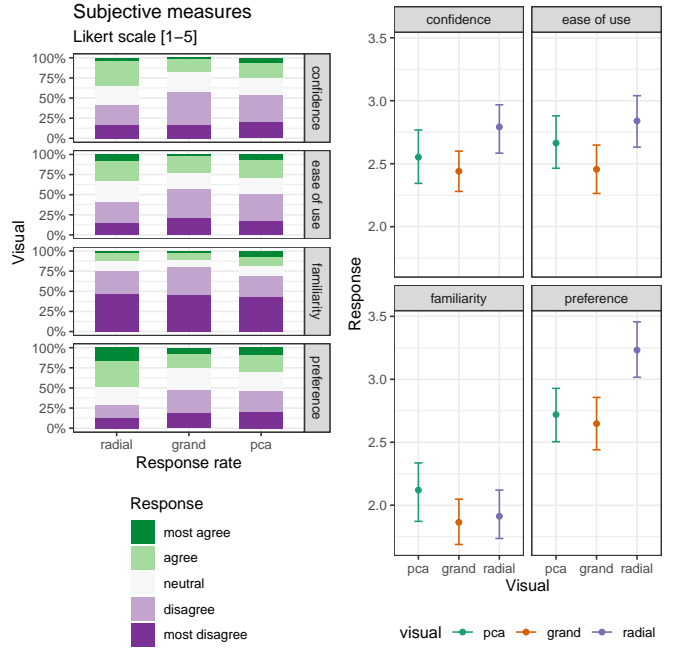

Fig. 8. The subjective measures of the 84 responses of the post-study survey with five-point Likert items levels of agreement. (L) Likert plots (stacked percent bar plots) with (R) mean and 95% CI of the same measures. Participants are more confident using the radial tour and find it easier to use than the grand tour. The radial tour is the most preferred visual.

subjective measures. The following were used in this study: confidence, ease of use, prior familiarity, and preference. Each of these questions was asked of all for each visual as 5-point Likert items.

The 84 evaluations of the post-study survey are shown in Fig. 8. The figure uses Likert plots or stacked percentage bar plots and asscoisated mean and 95% confidence intervals.

There was strong evidence to support that participants preferred the radial tour to either alternative. There is less evidence that the radial tour led to more confidence and was found easier to use than the grand tour. In confirmation of expectations, crowdsourced participants had low familiarity with all visuals, with no difference in mean supported.

5 DISCUSSION

Data visualization is an integral part of understanding relationships in data and how models are fitted. When it comes to multivariate data giving a comprehensive view quickly becomes difficult as the dimensions become sizable. Analysts have the task of choosing which visualization technique to use. Because the viewing volume/time of multivariate spaces typically increase quadratically with dimensions dimension reduction must be properly conducted. While there are optimization methods for static and animated visuals, the particular function used is a guided choice of the analyst.

Sect. 2 discussed various types of visualization which are may be preferred for differing tasks and ends. The visualization and perception of multivariate spaces is a broad and heterogeneous task. This work focuses a subset of linear projections and especially sheds light on potential benefit of providing user control in conjunction with the animated projection over many bases as a radial tour.

The radial tour is a method for the analyst to choose a variable to alter its contribution to the basis. The animation over small changes to the basis allows the sensitivity of the structure to be assessed from the variable contribution. The hypothesis is that user control over the basis and the permanence of observations between intermediate frames may lead to a better perception of the variable attribution causing the separation of clusters.

A mixed modeling analysis of the study provides strong support for this conclusion. That is, there is significant evidence to suggest the use of the radial tour leads to a sizable increase in accuracy. One unexpected caveat is that mixing the location of the signal at 33/66% almost completely negates this gain. Perhaps this is because the "half-clock" basis used did not give enough weight to the variable containing the small fraction. It was also interesting to note that no level of the experimental factors alone had a significant effect on this setup. Lastly, the follow-up survey asked participants to evaluate measures of the visuals. Most notably, participants preferred the radial tour to the other visuals. Knowing that the radial tour outperforms alternatives and is the preferred choice can help inform the selection of visual methods for developers and analysts.

There are several implicit limitations to this study: the task, type of data, and levels of the factors to name a few. The expansion of any of these areas is conceptually simple, but exponentially increases the number of participants needed to properly power the study. Additionally the sample of crowd-sourced, educated, but unfamiliar users may not extrapolate well to more experienced users. There are several ways that future work could be extended. Aside from expanding the support of the experimental factors, more exciting directions include: introducing a new task, including more visualizations, and changing the experience level of the target population. It is difficult to achieve good coverage given the number of possible factors to vary.

6 CONCLUSION

This paper discussed a crowdsourced mixed design user study ($n = 108$) comparing the efficacy of three linear projection techniques: PCA, grand tour, and radial tour. The participants performed a supervised cluster task, explicitly identifying which variables contribute to the separation of two target clusters. This was evaluated evenly over four experimental factors. In summary, mixed model regression finds strong evidence that using the radial tour sizably increases accuracy, especially when cluster separation location is not mixed at 33/66%. The effect sizes on accuracy are large relative to the change from the other

experimental factors and the random effect of data simulation, though smaller than the random effect of the participant. The radial tour was the most preferred of the three visuals.

There is no panacea for the comprehensive visualization of multivariate spaces. We have demonstrated that there is a definite value of user-control in linear projections. The agency of the analyst remains an important tool for the exploratory analysis of multivariate data.

ACKNOWLEDGMENTS

REFERENCES

[1] A. Adadi and M. Berrada. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE access*, 6:52138–52160, 2018.

[2] F. J. Anscombe. Graphs in Statistical Analysis. *The American Statistician*, 27(1):17–21, 1973. doi: 10.2307/2682899

[3] A. B. Arrieta, N. Díaz-Rodriguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, and R. Benjamins. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58:82–115, 2020.

[4] D. Asimov. The Grand Tour: a Tool for Viewing Multidimensional Data. *SIAM journal on scientific and statistical computing*, 6(1):128–143, 1985. doi: 10.1137/0906011

[5] P. Biecek. DALEX: explainers for complex predictive models in R. *The Journal of Machine Learning Research*, 19(1):3245–3249, 2018.

[6] P. Biecek and T. Burzykowski. *Explanatory Model Analysis: Explore, Explain, and Examine Predictive Models*. CRC Press, 2021.

[7] N. Cao, Y.-R. Lin, D. Gotz, and F. Du. Z-Glyph: Visualizing outliers in multivariate data. *Information Visualization*, 17(1):22–40, Jan. 2018. doi: 10.1177/1473871616686635

[8] R. B. Cattell. The scree test for the number of factors. *Multivariate behavioral research*, 1(2):245–276, 1966.

[9] J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Tukey. *Graphical methods for data analysis*. Chapman and Hall/CRC, 2018.

[10] C. Chang, T. Dwyer, and K. Marriott. An evaluation of perceptually complementary views for multivariate data. In *2018 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 195–204. IEEE, 2018.

[11] D. Cook and A. Buja. Manual Controls for High-Dimensional Data Projections. *Journal of Computational and Graphical Statistics*, 6(4):464–480, 1997. doi: 10.2307/1390747

[12] D. Cook, A. Buja, and J. Cabrera. Projection Pursuit Indexes Based on Orthonormal Function Expansions. *Journal of Computational and Graphical Statistics*, 2(3):225–250, Sept. 1993. doi: 10.1080/10618600.1993.10474610

[13] D. Cook, A. Buja, E.-K. Lee, and H. Wickham. Grand Tours, Projection Pursuit Guided Tours, and Manual Controls. In *Handbook of Data Visualization*, pp. 295–314. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. doi: 10.1007/978-3-540-33037-0_13

[14] D. J. Donnell, A. Buja, and W. Stuetzle. Analysis of Additive Dependencies and Concurvities Using Smallest Additive Principal Components. *The Annals of Statistics*, 22(4):1635 – 1668, 1994.

[15] N. Elmqvist, P. Dragicevic, and J.-D. Fekete. Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation. *IEEE transactions on Visualization and Computer Graphics*, 14(6):1539–1148, 2008. Publisher: IEEE.

[16] M. Espadoto, R. M. Martins, A. Kerren, N. S. T. Hirata, and A. C. Telea. Toward a Quantitative Survey of Dimension Reduction Techniques. *IEEE Transactions on Visualization and Computer Graphics*, 27(3):2153–2173, Mar. 2021. doi: 10.1109/TVCG.2019.2944182

[17] K. R. Gabriel. The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, 58(3):453–467, 1971.

[18] A. Gracia, S. Gonzàlez, V. Robles, E. Menasalvas, and T. von Landesberger. New Insights into the Suitability of the Third Dimension for Visualizing Multivariate/Multidimensional Data: A Study Based on Loss

[19] of Quality Quantification. *Information Visualization*, 15(1):3–30, 2016. doi: 10.1177/1473871614556393

[19] G. Grinstein, M. Trutschl, and U. Cvek. High-Dimensional Visualizations. p. 14, 2002.

[20] E. Kandogan. Star Coordinates: A Multi-dimensional Visualization Technique with Uniform Treatment of Dimensions. In *In Proceedings of the IEEE Information Visualization Symposium, Late Breaking Hot Topics*, pp. 9–12, 2000.

[21] W. Karwowski. *International Encyclopedia of Ergonomics and Human Factors, -3 Volume Set*. CRC Press, 2006.

[22] S. Lee, D. Cook, N. da Silva, U. Laa, N. Spyrison, E. Wang, and H. S. Zhang. The state-of-the-art on tours for dynamic visualization of high-dimensional data. *WIREs Computational Statistics*, n/a(n/a):e1573, 2021. doi: 10.1002/wics.1573

[23] D. J. Lehmann and H. Theisel. Orthographic star coordinates. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2615–2624, 2013. Publisher: IEEE.

[24] D. J. Lehmann and H. Theisel. Optimal Sets of Projections of High-Dimensional Data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):609–618, Jan. 2016. doi: 10.1109/TVCG.2015.2467132

[25] S. Liu, D. Maljovec, B. Wang, P.-T. Bremer, and V. Pascucci. Visualizing High-Dimensional Data: Advances in the Past Decade. *IEEE Transactions on Visualization and Computer Graphics*, 23(3):1249–1268, 2017. doi: 10.1109/TVCG.2016.2640960

[26] J. Matejka and G. Fitzmaurice. Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, pp. 1290–1294. ACM Press, Denver, Colorado, USA, 2017. doi: 10.1145/3025453.3025912

[27] J. A. McDonald. Interactive graphics for data analysis. Technical report, 1982.

[28] L. McInnes, J. Healy, and J. Melville. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv:1802.03426 [cs, stat]*, Dec. 2018. arXiv: 1802.03426.

[29] T. Munzner. Visualization analysis and design, 2014.

[30] J. E. Nam and K. Mueller. Tripadvisor^{ND}: A tourism-inspired high-dimensional space exploration framework with overview and detail. *IEEE transactions on visualization and computer graphics*, 19(2):291–305, 2012. Publisher: IEEE.

[31] L. Nelson, D. Cook, and C. Cruz-Neira. Xgobi vs the c2: Results of an experiment comparing data visualization in a 3-d immersive virtual reality environment with a 2-d workstation display. *Computational Statistics*, 14(1):39–51, Dec. 1999.

[32] M. d. Ocagne. *Coordonne'es paralle'les et axiales. Me'thode de transformation ge'ome'trique et proce'de' nouveau de calcul graphique de'duits de la conside'ration des coordonne'es paralle'les*. Gauthier-Villars, Paris, 1885.

[33] S. Palan and C. Schitter. Prolific: A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17:22–27, 2018.

[34] K. Pearson. LIII. On Lines and Planes of Closest fit to Systems of Points in Space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, 1901.

[35] T. L. Pedersen and D. Robinson. *gganimate: A Grammar of Animated Graphics*. 2020.

[36] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020.

[37] A. Sanchez, L. Raya, M. A. Mohedano-Munoz, and M. Rubio-Sanchez. Feature selection based on star coordinates plots associated with eigenvalue problems. *The Visual Computer*, 37(2):203–216, Feb. 2021. doi: 10.1007/s00371-020-01793-w

[38] A. Sanchez, C. Soguero-Ruiz, I. Mora-Jimenez, F. J. Rivas-Flores, D. J. Lehmann, and M. Rubio-Sanchez. Scaled radial axes for interactive visual feature selection: A case study for analyzing chronic conditions. *Expert Systems with Applications*, 100:182–196, June 2018. doi: 10.1016/j.eswa.2018.01.054

[39] M. Sedlmair, T. Munzner, and M. Tory. Empirical Guidance on Scatterplot and Dimension Reduction Technique Choices. *IEEE Transactions on Visualization & Computer Graphics*, (12):2634–2643, 2013.

[40] C. Sievert. *Interactive Web-Based Data Visualization with R, plotly, and shiny*. Chapman and Hall/CRC, 2020.

[41] J. P. Snyder. *Map projections–A working manual*, vol. 1395. US Government Printing Office, 1987.

[42] N. Spyrison and D. Cook. spinifex: an R Package for Creating a Manual

Tour of Low-dimensional Projections of Multivariate Data. *The R Journal*, 12(1):243, 2020. doi: 10.32614/RJ-2020-027

[43] J. W. Tukey. *Exploratory Data Analysis*, vol. 32. Pearson, 1977.

[44] L. van der Maaten and G. Hinton. Visualizing Data Using t-SNE. *Journal of machine learning research*, 9(Nov):2579–2605, 2008.

[45] J. A. Wagner Filho, M. F. Rey, C. M. Freitas, and L. Nedel. Immersive visualization of abstract information: An evaluation on dimensionally-reduced data scatterplots. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 483–490. IEEE, 2018.

[46] Y. Wang, J. Li, F. Nie, H. Theisel, M. Gong, and D. J. Lehmann. Linear Discriminative Star Coordinates for Exploring Class and Cluster Separation of High Dimensional Data. *Computer Graphics Forum*, 36(3):401–410, 2017. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13197. doi: 10.1111/cgf.13197

[47] H. Wickham, D. Cook, and H. Hofmann. Visualizing statistical models: Removing the blindfold. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 8(4):203–225, 2015. doi: 10.1002/sam.11271

[48] H. Wickham, D. Cook, H. Hofmann, and A. Buja. tourr: An R Package for Exploring Multivariate Data with Projections. *Journal of Statistical Software*, 40(2), 2011. doi: 10.18637/jss.v040.i02

[49] Y. Xie, J. J. Allaire, and G. Grolemund. *R Markdown: The Definitive Guide*. Chapman and Hall/CRC, Boca Raton, Florida, 2018.

[50] I. Yanai and M. Lercher. A hypothesis is a liability. *Genome Biology*, 21(1):231, 2020. doi: 10.1186/s13059-020-02133-w