SPLOCIS – Extending SPLOMs to a Scatterplot Cube with Interactable Shadows for Immersive Analysis in Virtual Reality

Melanie Derksen* TU Dortmund University

Viktoria Diekel† TU Dortmund University

Torsten Kuhlen‡ RWTH Aachen University

Mario Botsch § TU Dortmund University

Tim Weissker¶ RWTH Aachen University

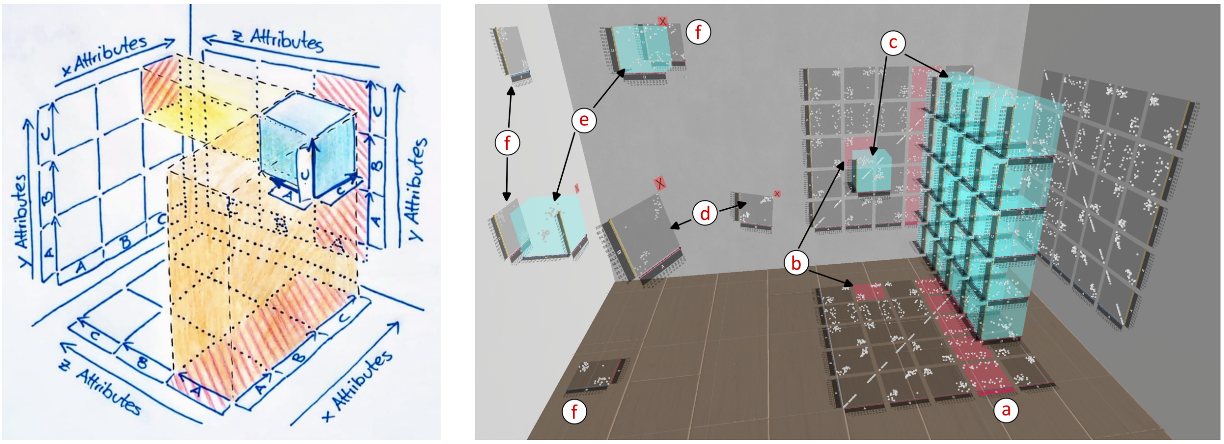

Figure 1: We present SPLOCIS – a new concept for investigating 2D and 3D scatterplots in an immersive virtual environment. While scatterplot matrices (SPLOMs) showcase every attribute combination in 2D plots arranged in a matrix, the concept of SPLOCIS covers all possible 3D plots arranged in a cube. Two walls and the floor are decked with interactable SPLOMs, that represent the same data in a transformed manner. *Left*: As showing all 3D plots at once would lead to visual clutter, they are invisible per default. They can be turned visible by selecting the labels located at the borders of the SPLOMs (for the z-axis attribute A is selected) and/or by selecting individual 2D plots in the SPLOMs directly (the 2D plot with C on its x- and y-axis is selected). This way, the user specifies step-by-step which attributes are of interest, filtering which 3D plots should remain visible for further investigation (the 3D plot with attribute C on its x- and y-axis and A on its z-axis is visible). *Right*: In the final application (here, the label (**a**), and 2D plots selections (**b**) result in a subset of visible 3D plots (**c**)) users can extract plots from the SPLOCIS, creating independent, transformable plot instances (**d**, **e**) for a more in-depth analysis of the data. 3D plots cast projections on close walls if they are parallel to them (**e**) in the form of 2D plots (**f**), enabling a dynamic and interactive data exploration experience.

Abstract

In data analysis, scatterplots serve as an initial tool for exploring the relationships between two or three attributes. While scatterplot matrices (SPLOMs) display every attribute combination through numerous 2D scatterplots to show a concise overview of a multivariate dataset, this approach is not directly suitable for 3D scatterplots due to visual clutter. Since research has shown that immersive virtual environments can enhance data analysis compared to traditional 2D desktop setups – especially for spatial analysis tasks – we propose an interactive system, called SPLOCIS, that makes use of virtual reality to enable users to interactively filter and select 3D scatterplots from all possible attribute combinations. Our user study, combining both qualitative and quantitative results, demonstrates that SPLOCIS is a particularly novel and stimulating approach to work with multivariate data in immersive environments. It enables solving classic data exploration tasks in an efficient and accurate way, while not imposing unexpectedly high task loads. Moreover, our findings provide promising suggestions for further developments.

Index Terms: Virtual reality, 3D user interfaces, Head-mounted display, Immersive analytics, Scatterplot, Scatterplot matrix.

1 Introduction

As data collection becomes increasingly automated across nearly all domains, the resulting high-dimensional and complex datasets are growing rapidly, thereby increasing the demand for effective visualization and analytical tools [41]. Well-designed visual representations are often more intuitive and effective at revealing insights than standard statistical methods [8]. When beginning a data analysis process, scatterplots provide a solid visual basis for initial visual exploration. They help examining unusual patterns, finding correlations [24], identifying clusters and outliers, and support overall data understanding [28]. When dealing with multivariate data, scatterplot matrices (SPLOMs) extend this capability by displaying all pairwise combinations of variables in 2D plots. They provide a quick overview of potential relationships between attributes and characterize their nature, reveal outliers, and identify any clustering patterns within the data based on groups [21, 28].

Previous research demonstrates the suitability of virtual reality (VR) for data analysis, highlighting its advantages over traditional 2D desktop setups particularly for spatial analysis tasks [19, 35, 59].

*e-mail: melanie.derksen@tu-dortmund.de
†e-mail: viktoria.diekel@tu-dortmund.de
‡e-mail: kuhlen@vr.rwth-aachen.de
§e-mail: mario.botsch@tu-dortmund.de
¶e-mail: me@tim-weissker.de

However, identifying and retrieving the most relevant plots from a large multivariate dataset still remains an open challenge in real-life data analysis scenarios. Even though the aforementioned SPLOMs are suitable for gaining an overview of the data and help data analysts in their exploration process (so which plot they intend to investigate in greater detail), their two-dimensional nature of attribute presentation is optimized for conventional desktop displays. While it is technically possible to display them on a 2D plane within an immersive virtual environment (iVE), the benefits of this analysis approach over a desktop display are questionable. However, the additional dimension and immersive nature of iVEs offer novel opportunities for extending the idea of SPLOMs into 3D, thereby revealing three-way attribute relationships in the data that cannot be explored with SPLOMs alone.

Therefore, this work aims to deal with the overarching research question of **how to make 3D scatterplots effectively, comfortably and intuitively explorable in iVEs**. To do so, we introduce and evaluate a novel immersive data exploration approach, called "Scatterplot Cube with Interactable Shadows" (SPLOCIS). Our concept, guided by well-known best practices from HCI [45, 54], is based on traditional SPLOMs which form the "interactable shadows" (IS) and serve as interaction and information basis (see Figure 1). Moreover, it enables to explore and visualize all triple-combinations of attributes by extending the concept of SPLOMs by a third spatial dimension. In order to prevent visual clutter, all 3D plots are invisible per default – the user is enabled to specify certain sets of attributes per axis, making only those 3D plots visible that match/fulfill their individual conditions. In a user study, we evaluate how typical scatterplot tasks affect performance and cognitive load within our application (RQ1). Furthermore, we evaluate the intuitiveness, accessibility, and suitability of our suggested interaction metaphors for 3D data exploration (RQ2). As our system consists of three SPLOMs in order to "span the 3D plot space", users are confronted with plots in which two or more axes display identical attributes, similar to the diagonal in classic SPLOMs. Therefore, we explore the question whether these plots are either disturbing or supporting the exploration process (RQ3). In summary, our research led to the following contributions:

- the introduction of SPLOCIS, which is a novel interactive method for investigating high dimensional data in an iVE,
- quantitative results of a user study with 33 participants, showing that our system provides a novel, stimulating, efficient, and attractive way to solve typical scatterplot exploration tasks in higher dimensions, while not inducing unexpectedly high task loads compared to other systems in the literature,
- qualitative results based on self-reports, emphasizing the application's fun factor, intuitiveness and understandability, and revealing promising areas for further developments.

2 Related Work

We begin by reviewing the general strengths and limitations of scatterplots and scatterplot matrices (SPLOMs) (Section 2.1), followed by a review of existing methods that address their known drawbacks. We then examine the advantages of a third spatial dimension and immersive data exploration (Section 2.2), motivating the rationale behind our design decisions. Lastly, we survey existing immersive systems from other domains to highlight the practical potential of iVEs for data-related work (Section 2.3).

2.1 Scatterplots and SPLOMs

Given their effectiveness and widespread use, it is not surprising that scatterplots and SPLOMs are the focus of extensive research [24, 29, 55].

Scatterplots are one of the most used techniques in data visualization [55]. Showing all of the data points, they are a powerful statistical tool for analyzing the relationship between two measured or observed variables [8]. Compared to alternatives for information visualization, such as parallel coordinate plots (PCPs), scatterplots have often been validated as the most effective solution for tasks such as identifying pairwise relationships, tracing values across dimensions, and outlier identification [9, 16, 17, 24, 29, 33, 36, 46]. However, most datasets consist of more than two variables, making it necessary to explore all potential combinations of variables. This is where SPLOMs come into play. Following the concept of small multiples [45, 56], they provide a grid of scatterplots, containing all combinations of variables. By doing so, they allow for a holistic view of multiple variable relationships as well as detailed insights at once, making them a powerful tool for exploratory data analysis, and are shown to be faster and more accurate for outlier detection and correlation estimation than PCPs [9].

Despite their benefits, scatterplots and SPLOMs are associated with numerous challenges, such as handling a large number of data points or representing highly multidimensional data [7, 14, 42, 51, 62, 63]. Those challenges have also already been widely addressed in prior research. E.g., various techniques have been developed to mitigate overplotting in SPLOMs, including the use of grayscale or symbol area to represent density [7], as well as splatterplots [42], which represent dense regions as smooth, closed shapes while explicitly highlighting representative outliers. In case of a large number of dimensions within the dataset, there are approaches like Glyph SPLOMs, a technique to summarize a SPLOM [63]. Dang and Wilkonson [14] introduce ScagExplorer, an interactive platform designed to explore large spaces of scatterplots by visualizing scagnostics – quantitative measures that characterize the shape of 2D distributions [60, 61]. Rather than manually inspecting thousands of scatterplots in high-dimensional datasets, users can leverage scagnostic measures to identify patterns and detect anomalies. This enables more efficient analysis of SPLOMs that would otherwise be too large to interpret directly. The work of Sarikaya et al. [51] addresses the challenges of designing effective scatterplots for complex, large-scale data by surveying analysis tasks, data characteristics, and existing scatterplot design variations. It connects these elements to guide design decisions and offers best practices, open questions, and challenges to support practitioners in selecting suitable scatterplot designs for specific analytical goals. Yang et al. [62] present the Value and Relation (VaR) display, a multidimensional visualization technique designed to minimize visual clutter and support interactive exploration of high-dimensional data. By making relationships among dimensions visible, the VaR display supports efficient navigation, reveals patterns, and enables detailed analysis even in datasets with hundreds of dimensions.

This research shows that there are many suggestions and approaches to overcome the limits and drawbacks of SPLOMs and scatterplots. A range of techniques, such as splatterplots, glyph-based summaries, scagnostics, and advanced design frameworks, have been developed to address these issues and enhance the interpretability and scalability of scatterplot and SPLOM visualizations. Based on their simplicity as well as the validated advantages of scatterplots and SPLOMs in previous work, we decided to use them as a basis for our developments in this paper.

2.2 Benefits of 3D and Immersive Data Exploration

While scatterplots aid in examining paired data values, numerous data analysis scenarios involve multivariate data. Just as bivariate data is represented in a two-dimensional coordinate system, trivariate data can be plotted in a three-dimensional coordinate system [8].

Gracia et al. [22] compare scatterplot representations of multidimensional multivariate data reduced to two and three dimensions, respectively. They found that the errors made by participants are smaller in their 3D visualizations, especially in tasks involving distance assessment. While their participants complete tasks faster in 2D, the drop in quality is significant when moving from 3D to 2D, highlighting the value of the third dimension for more precise data interpretation. On top of that, Poco et al. [49] confirm that 3D projections outperform 2D projections in terms of precision. Their user study indicates that some cluster-related tasks can be more reliably and confidently answered with 3D projections.

Furthermore, the findings of McIntire and Liggett [43] suggest that stereoscopic 3D visualizations can offer notable advantages for tasks such as precise spatial localization of objects and manual interaction with data or virtual information, and can yield performance benefits aligned with cognitive advantages, leading to an increased understanding of spatial and/or multidimensional data.

Etemadpour et al. [19] find that stereoscopic virtual environments can enhance the perception of distances in 3D visualizations, leading to better performance in local analysis tasks compared to traditional 2D screens. On top of that, Wagner Filho et al. [59] find that using head-mounted displays (HMDs) for exploring 3D scatterplots can reduce the effort in finding information and enhance users' perceived accuracy and engagement compared to desktop setups.

Kraus et al. [35] examine how immersion affects cluster identification in abstract scatterplot visualizations. Their results show that 3D visualizations, especially in their iVE, outperform 2D screens in terms of accuracy, efficiency, orientation, memorability, and user preference, concluding that higher immersion can be a substantial benefit in the context of 3D data exploration and cluster detection.

Besides that, Allcoat and von Mühlenen [3] infer that VR can be especially beneficial for data interpretation tasks, can lead to a high engagement, can increase positive and decrease negative emotions.

Collectively, these findings suggest that a third spatial dimension and iVEs can offer significant advantages regarding data exploration compared to traditional 2D data representations. Therefore, we decided to focus our work on improving the exploration of 3D scatterplots in iVEs.

2.3 Existing Tools for Immersive Data Exploration

Motivated by the benefits of adding a third spatial dimension and moving to iVEs, numerous tools and interaction metaphors have been proposed to assess the practicability of iVEs for data analysis.

Hurter et al. [32] propose FiberClay, a multidimensional visualization system for interactive 3D trajectory analysis in iVEs that leads to a better understanding of dense and complex datasets. Hayatpur et al. [27] present DataHop, a novel visualization system that enables users to lay out their data analysis steps in an iVE. Their user study suggests that spatially mapping workflows (by assigning distinct spatial positions to data filtering steps) can support exploration and understanding of multidimensional datasets.

Wagner Filho et al. [58] evaluate VirtualDesk, where users explore immersive data, rendered within arm's reach, in a seated position. Their study shows that their system provides high comfort and immersion, matches or even outperforms its desktop-based equivalent in all analytical tasks, with minimal to no additional time overhead, and enhances users' perceived efficiency and engagement.

Cordeil et al. [13] introduce ImAxes, an immersive system for exploring multivariate data, in which users can interact directly with virtual data axes like physical objects and combine them into sophisticated visualizations. Batch et al. [5] gather requirements and enhance the ImAxes prototype to support professional economic analysis. Surprisingly, they discovered that well-known limitations of VR, such as fatigue or text legibility, are not of significant concern in their study setup. Participants, including those new to VR, are overwhelmingly positive about their experience. They report a higher level of engagement and presence within the iVE compared to traditional desktop setups and found that creating visualizations was faster and easier. Although ImAxes supports various types of data visualizations, getting an overview of attribute correlations requires users to manually create a SPLOM, a process involving many operational steps and significant time investment. Moreover, ImAxes does not support selectively displaying plots within a 3D SPLOM, which can lead to visual clutter. It also does not allow to extract specific plots from a 2D or 3D SPLOM – instead, it requires a manual recreation of the plots for detailed investigations.

Le et al. [38] introduce FIESTA, a system that enables users to collaborate flexibly untethered from physical display devices. Unlike many existing systems, FIESTA supports co-located collaboration in a shared room-sized environment, where users can create, position, and interact with immersive data visualizations.

While Section 2.2 focused more on abstract task types and the general applicability of a third dimension and immersive scenarios compared to less immersive or lower-dimensional setups, the works and findings discussed here highlight the practical applicability of iVEs for data-related work. In sum, the applications presented here are generally well accepted, help increase data understanding and user engagement, even among users who are less familiar with VR.

Overall, the proven effectiveness of scatterplots and SPLOMs for data exploration, combined with the advantages of a third spatial dimension and the success of existing immersive applications, motivate our approach to develop an immersive and interactive exploration system based on scatterplots and SPLOMs. While ImAxes already enables immersive data interaction and exploration, it requires many manual steps to gain an overview of the dataset. In contrast, our approach aims to reduce the number of required interactions while supporting both an overview of 2D and 3D plots as well as the deeper investigation of specific plots.

3 Data Exploration with SPLOCIS

In this section, we present our novel VR system called SPLOCIS – "Scatterplot Cube with Interactable Shadows", designed for interactively and efficiently investigating 2D and 3D scatterplots by extending the idea of classic SPLOMs. For that purpose we used the Unity game engine in combination with the Immersive Analytics Toolkit (IATK) [12]. This integration allows for rapid prototyping of data visualization systems within VEs. While we specialized our developments on the *Meta Quest 3* [1] and its dedicated controllers, the underlying concepts can be applied to other HMDs as well.

To systematically approach the task to make 3D scatterplots explorable in an iVE, benefiting from the immersive facilities while not overloading the virtual space with plots and information, we defined three design requirements that guided our developments:

R1: Users should easily get an overview of not only 2D but also 3D data relationships.

R2: The visualization should reduce visual clutter caused by too many plots displayed at a time as much as possible.

R3: The interaction design should follow established best practices from HCI and information visualization research.

Physical Setup We focused on a stationary, seated VR experience, motivated by typical working conditions of data analysts who perform their tasks usually in a desk-based office environment with limited room for movement. Additionally, previous research has shown that seated VR experiences tend to cause less fatigue than standing experiences [10] and are also associated with reduced levels of cybersickness [44, 50].

Virtual Setup In our system, data analysis takes place in a minimalist virtual room featuring white walls and one SPLOCIS instance that contains one preselected dataset (see Figure 1, *right*). As additional attributes require a larger space, and past work could show that smaller spaces were found to be significantly less attractive, efficient, and stimulating when working with large amounts of data [40], the room dynamically scales with the number of attributes displayed in the SPLOCIS, with the walls always being more than twice as long as their respective SPLOM (while each plot within SPLOCIS has an default axis length of 0.3 m). This ensures sufficient space for data visualization and interaction.

As the application setup is physically stationary, but the amount of data displayed in the iVE might require navigation to be fully explorable, we made use of a standard rotation and translation snapping implementation provided by Unity's *XRInteraction Toolkit* [2], as those techniques are known to significantly decrease symptoms of cybersickness and nausea levels, especially with longer VR exposure time [20]. Thus, according to the XRInteraction Toolkit's default values users are enabled to rotate left/right by 45° by pushing the thumbstick of the right controller left/right and rotate by 180° by pulling the thumbstick down. Translation can be done by pointing at an arbitrary position on the virtual floor while not exceeding a distance of 10 m; a pictorial projectile on the floor as well as a ray connecting the projectile and the respective virtual controller representation indicates the teleportation goal.

Together with the physical setup, those design decisions are supposed to build a good prerequisite for even longer working sessions within our VR system.

Building a Scatterplot Cube The main idea of SPLOCIS lies in extending the concept SPLOMs by a third spatial dimension and transferring it into an immersive space, resulting in what we call a scatterplot cube (SPLOC). While a SPLOM displays all pairwise combinations of attributes using 2D plots arranged in a matrix, the SPLOC expands this idea by incorporating all possible triple combinations of attributes into 3D scatterplots arranged in a cube (R1).

Unfortunately, visualizing multiple 3D plots and arranging them that densely can result in visual clutter, which can make it difficult for users to perceive structure and understand the content of the visualization [47]. To mitigate this issue (R2), all plots of the SPLOC remain hidden by default, and it requires user interaction to reveal individual plots. For this, an effective interaction method (R3) that allows users to selectively display relevant 3D plots without cluttering the scene (R2) was needed. Our goal, however, is not only the possibility to explore 3D plots, but to support both 2D and 3D data exploration (R1) at the same time. We address both of these issues by integrating three classic SPLOMs. Given that SPLOMs are a well-established visualization technique in 2D, and since the concept of the SPLOC builds directly upon them, they should be a natural approach to reach both of those goals in our system.

Each SPLOM is therefore an interactable shadow (IS) and can be seen as a 2D projection of the SPLOC, positioned on one out of three spatial planes, being the room's walls for the x-y- and y-z-plane and the floor for the x-z-plane (see Figure 1). Depending on the preconfigured attributes per room axis, those SPLOMs may contain the same information about the data in a transformed way. These SPLOMs serve two purposes: first, they provide an overview of the dataset through 2D attribute relationships (R1), embodying the "Overview First" principle, which implies to show the entire dataset with a display method that works best in the individual case [54] (R3); second, they act as an interface for controlling the visibility of 3D plots within the SPLOC, reducing visual clutter on the one hand (R2) and following the concept of "Zoom and Filter" on the other hand, which means to remove extraneous data and give the data of interest more detail [54] (R3). The combination of SPLOMs and the SPLOC also follow the principle of small multiples [45, 56], providing both an overview (R1) and detailed insights, which should make our system particularly valuable in exploratory data analysis without missing important information [57] and without overloading visual working memory [48] (R3). Users can control the visibility of 3D plots by setting conditions the plots have to fulfill to turn visible. In this context, "conditions" mean the specification which attributes the 3D plots should have on which spatial axis. The user can either specify one attribute for one axis by selecting a label at the border of a SPLOM (see Figure 1, *right*: **a**) or specify two attributes for two axes at a time by selecting a 2D plot inside a SPLOM (see Figure 1, *right*: **b**). The latter implicitly defines two conditions, as each 2D plot maps two attributes to two spatial axes. This stepwise displaying and filtering process can be seen as calculating volumetric intersections. Our supplementary material provides the mathematical formalization of this process and a video, in which the procedure is shown with examples.

Plot Clones While plots within the SPLOMs and SPLOC are static regarding their transformations, users can create transformable and scalable 2D and 3D plot clones from them (see Figure 1, *right*: **d**, **e**). This way, we follow the concept of "Details on Demand", which means to let users reveal more details of the data only when needed [54] (R3). This cloning should enable a more in-depth exploration of individual plots, independent of the broader dataset within the "original" SPLOCIS instance. It should also allow for a flexible arrangement of data items within the virtual space, which is expected to facilitate a more intuitive sense-making process [39] and past work could show that users tend to make use of rearranging the data items they work with [52].

If a 3D plot clone is dropped nearly parallel (meaning it requires at most a 17° rotation to become parallel) to at least one of the six virtual room's boundary planes (being the walls, ceiling, and the floor) it "snaps", so that the faces of the plots are parallel to the opposing plane. Additionally, if the plot is also close enough to a plane (less than 1.5 m), then a 2D projection of it is cast on it (see Figure 1, *right*: **f**). This projection is therefore again a 2D scatterplot the user can interact with, and includes the same two attributes that are on the 3D plot side facing the respective plane.

4 User Study

We designed and conducted a user study to systematically evaluate our proposed SPLOCIS system. Our goal was to approach our global research question of how to design an immersive system that supports an efficient, comfortable, and intuitive exploration of multivariate data relationships. To do so, we broke this high-level question down into more precise low-level research questions, which we will explain in the following and aim to answer with our user study.

4.1 Research Questions

Past work could show that especially spatial analysis tasks benefit from the immersive qualities of VR. Therefore, we aim to identify which task types work well in our system and which do not, guiding further developments of VR data exploration tools to reduce cognitive load while preserving performance.

RQ1: How does the task type in the application affect performance and cognitive load? Are there task types that are particularly (un)suitable?

In order to support data analysts, an immersive data exploration tool must be not only functional but also accessible and easy to learn. This is why we explore our participants' subjective impressions and gather feedback to identify strengths, weaknesses, and opportunities for improving immersive data exploration.

RQ2: How intuitive/accessible/suitable is the application for working with multivariate data by exploring 3D scatterplots? Which features would make the tool (or data analysis in VR in general) better?

Since one SPLOM typically contains repeated axis assignments (being a plot with twice the same attribute) on one of its diagonals, it is apparent that incorporating three SPLOMs to span the 3D scatterplot space introduces even more repeated axis assignments and redundant data. On top of that, our main focus is on the exploration of 3D plots having not only two but three axis placeholders. Consequently, we need to consider that those plots can have either three distinct attributes or two (or even three) identical ones. While the first case is full of informative value, the latter cases contain redundancies and might be less information-rich during data exploration. On the other hand, multiple axes with the same attribute might also yield novel insights, similar to the histogram-like views on the diagonal of SPLOMs. Therefore, our objective is to investigate if repeated axis assignments are an issue, whether or not they support the data exploration experience or rather lead to confusion and distraction.

RQ3: To what extent do repeated axis assignments disturb or support the exploration process?

4.2 Apparatus

For the study, we used the *Meta Quest 3* having a resolution of 2064×2208 pixels per eye and an update rate of 120 Hz. The application is based on Unity (Version 2022.3.15f1). Participants had an interaction space of about 1 m × 1 m and sat on a swivel chair enabling 360° rotations, providing an unrestricted field of regard.

4.3 Procedure

Based on a list of criteria from both participating institutions, our study did not require an ethics approval, given that our study did not elicit strong emotions or stress, induce pain or harm, deceive participants, conceal risks, capture biological material, or pose security-relevant threats.

Participants were invited to our lab, were informed about the purpose of the study, agreed to participate voluntarily, and signed a consent form. To ensure a consistent level of understanding regarding scatterplots and SPLOMs across all participants, they went through some informative slides providing the necessary knowledge basis to solve the tasks in the study appropriately.

Afterwards, the participants were introduced to the VR system with a brief descriptive and interactive tutorial. Here, they were given sufficient time to practice the handling as well as to complete some exemplary tasks, which were of the same kind as in the actual study. To initially set the focus on learning the system and avoid overwhelming users with data, during the tutorial the SPLOCIS instance included four attributes, whereas it included five in the main study. Once participants were comfortable with the system, the actual study started, during which they went through the task types explained in Section 4.4. After completing a task set of one type, they filled out two questionnaires. First, the Raw-TLX [26, 25] to assess the perceived task load for the respective task type to check whether or not our participants could focus on data exploration or if they were overwhelmed using the system. Second, the Fast Motion Sickness Scale (FMS) [34] to measure symptoms of cybersickness. Since cybersickness is a general usability concern in VR we wanted to ensure that our system does not induce significant levels of it. Besides the subjective questionnaires during the VR part, we logged the task completion times and task correctness.

When the VR part was completed, the participants filled out the User Experience Questionnaire (UEQ) [37] as well as the System Usability Scale (SUS) [6] to comprehensively cover classic usability as well as user experience aspects. They then rated custom statements on the system's intuitiveness, accessibility, and suitability for data exploration. They also provided up to three positive and negative aspects of the system in free-text form, potential use cases they had in mind for repeated axis assignments, and which features would improve working with 3D scatterplots in immersive spaces. Lastly, we asked for self-assessment regarding previous knowledge about VR, gaming, scatterplots and data exploration in general, and finally captured some demographic data.

The entire procedure took approximately 90 min to complete and participants received an expense allowance of 18€.

4.4 Tasks and Datasets

During the VR part of the study, all tasks participants had to solve followed the same structure: The participants had to interact with a SPLOCIS instance in order to find a 3D plot matching specific properties that were asked for on a UI element in the virtual scene. This plot then had to be put into a box. During the tutorial, this box' appearance changed to green with a white check mark when the plot fulfilled the requirements, and to red with a white cross otherwise, to help participants develop a clearer understanding of which plots are correct and which are not before the actual study started. When it came to the actual study the box did not provide any feedback on the correctness so that participants were fully focused on the current task and not influenced by either positive or negative feedback regarding their performance in past tasks. The participants were told to work as quickly as possible during the actual study.

To avoid individual contextual reference and to have full control over features present in the data and therefore the number of correct as well as misleading plots, we created synthetic datasets each consisting of 40 data points and five attributes named A, B, C, D, E with values ranging from 0 to 100.

There were five different types of tasks participants had to solve and which were performed one after another. While the first task type tests basic interaction mechanics of the tool, the remaining types of tasks followed typical scatterplot/SPLOM analysis tasks, as mentioned by Sarikaya and Gleicher [51] as well as Heckert et al. [28]. Per type there were four tasks, which were performed successively. To ensure a similar level in difficulty for each of the four tasks per type, the number of plots fulfilling the requirements remained the same during one task type. We provide the datasets as well as task details in our supplementary material.

T1: Attributes For this task type, participants had to pick out plots having three specific attributes, whereby it did not matter on which axis which attribute was located.

T2: Correlations For this type, two kinds of tasks were possible: finding a plot with pairwise linear correlations (here, it did not matter whether they were positive or negative), or finding a plot with pairwise and explicitly positive linear correlations. In this context, "pairwise" means that every combination of attributes in a potentially correct plot had to have either a linear correlation in general (for the first case of task) or a positive one (for the latter case). To make sure our data contained recognizable linear correlations, we performed pairwise correlational analyses by computing Pearson's r, ensuring that potentially correct plots had a large correlation coefficient, meaning that $|r| > 0.5$.

T3: Clusters Here, the participants were asked to find plots having two clusters. To ensure the existence and verifiability of clusters we ran DBSCAN [18] on every 3D plot. A plot was considered as being a valid answer if DBSCAN found two clusters in it with the parameters $\varepsilon = 1.3$ and a minimum samples value of 9.

T4: Outlier For the outlier identification tasks, participants had to find a plot with one outlier. We ensured that plots having three different attributes were normally distributed by applying the Henze-Zirkler test for Multivariate Normality [30] with a significance level of 0.05. To prove that specific plots had at least one considerably visible outlier, we took the Mahalanobis distance [15] into account, which is a statistical measure to identify multivariate outliers. We selected a critical alpha level of 0.001 as it was a good balance that made outliers noticeable without making them stand out too prominently. The corresponding critical χ^2

value for three degrees of freedom (since we examine 3D plots) at our critical alpha level is 16.266. Thus, if the maximum squared Mahalanobis distance for the data distribution in one plot exceeded this critical χ^2 value, this plot had at least one proven outlier.

T5: No Relation This task type asked for plots having none of the aforementioned relationships. This means that there were neither correlations (meaning according to Pearson there were not even small correlations $|r| < 0.1$) nor multiple clusters found by DBSCAN with the aforementioned parameters.

To prevent participants from memorizing certain data patterns, the dataset switched after each task – except for the first type of tasks being the identification of plots by attributes. For this task type, the distribution does not affect the task's difficulty. Together with the fact that the attribute names remained the same in each dataset, switching the dataset would probably have had no effect.

4.5 Participants

33 participants (8 female, 24 male, and 1 preferred not to say) between 20 and 49 years of age ($M = 27.2$, $\sigma = 6.36$) attended the user study. They were recruited on the local university campus and via dedicated mailing lists. Self-reports on prior VR, data visualization, and scatterplot experiences are listed in Table 1.

Table 1: Total counts of self-reported prior experiences. The individual experience levels were measured on a five-point Likert scale, with 1 indicating "very little" and 5 "very much" experience; the last column shows the mean values and standard deviations, respectively.

Type of Experience	1	2	3	4	5	(M, σ)
VR	14	6	3	2	8	(2.52, 1.66)
Data Visualization	3	8	12	8	2	(2.94, 1.06)
Scatterplots	10	10	10	3	0	(2.18, 0.98)

5 Results

This section reports on the results of our data analysis, beginning with an overview of FMS scores as an indicator for the validity of all other measurements. Throughout this paper, confidence intervals (CIs) are reported at the 95% level.

5.1 Cybersickness

The FMS scores, gathered after each task type (resulting in a total of $N = 165$ scores; with a possible range from 0 to 20), ranged in our study from 0 to a maximal value of 13 ($M = 1.109, \sigma = 2.115$). The majority of scores ($N = 114$) were 0, indicating that participants most of the time felt no sickness at all. A smaller number of scores ($N = 49$) fell between 1 and 6. Quite high scores of 12 and 13 were reported only once each.

5.2 Task Load, Task Completion Time, and Correctness

The distributions of task load scores per task type measured with the Raw-TLX questionnaire [26] are shown in Figure 2. In order to evaluate the results, we use the values provided by Grier [23] and Hertzum [31], since the original publication of this questionnaire does not come along with benchmark values and our user study handled one single condition only, therefore having no apparent reference system for comparison.

Across all task categories, the mean values are notably lower than the averages reported in Grier's general dataset ($M = 45.29$, $\sigma = 14.99$ for the Raw-TLX), Hertzum's general dataset ($M = 42$, $\sigma = 13$), and Hertzum's specific VR dataset ($M = 41$, $\sigma = 15$). Additionally, most 75th percentiles of measured task loads are between 25 (T5: **No Relation**) and 36.67 (T2: **Correlations**), and therefore also below the averages reported by Grier and Hertzum. Only the 75th percentile of the **Outlier** task type (T4) exceeds the benchmark values with 48.33. Moreover, nearly all CIs are below the benchmark averages – here again, only the **Outlier** type (T4) builds an exception by slightly exceeding the average of Hertzum's specific VR dataset. On top of that, the **Outlier** task type (T4) features the overall maximal reported task load of 75, and is the only type which no participant rated with a task load of 0.0.

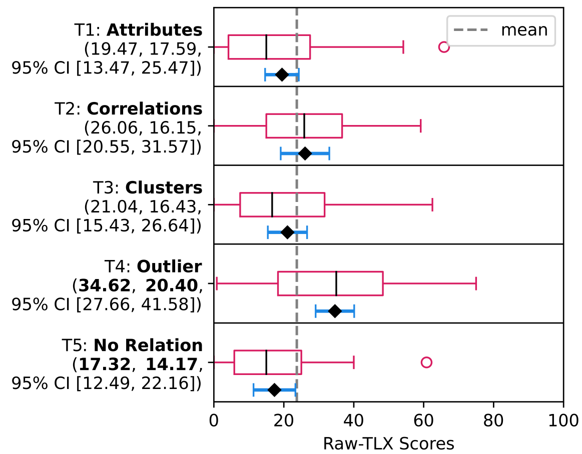

Figure 2: Boxplots illustrating the Raw-TLX values per task type. Below each boxplot a diamond represents the mean with bars that indicate the corresponding CIs. The first entry inside the brackets is the mean value for the respective type, the second one the standard deviation, followed by CIs. Minimal and maximal mean values and standard deviations are bold.

We further report (in Table 2) on the mean duration to complete a task, meaning the average time the participants needed to find an appropriate plot and put it into the box, and on the mean task correctness, meaning the average percentage of correctly chosen plots fulfilling the respective task requirements. A common mistake in T2 to T5 was to select a plot that did not have three distinct attributes, independent of the pattern asked for in the task. Taking the average correctness rate of all tasks into account, being 93.78% (see Table 2), this kind of mistake accounts for 2.88% correctness rate reduction, representing a total of 46.34% error causes.

Additionally, we observe a positive correlation between the task load and task completion time, $r = 0.45$, which can be considered medium based on the threshold values reported by Cohen [11]. We further observe small negative correlations between task completion time and correctness, $r = -0.19$, and task load and task correctness, $r = -0.15$.

5.3 User Experience and Usability

The UEQ aggregates its 26 individual responses into six higher level categories with scores between -3 and +3 to quantify different aspects of usability. The results for our system as well as comparisons with the benchmark dataset of the UEQ are shown in Figure 3. Most positive scores were obtained for the categories *Stimulation* ($M = 2.1$, $\sigma = 0.81$, 95% CI $[1.82, 2.37]$) and *Novelty* ($M = 1.96$, $\sigma = 0.83$, 95% CI $[1.68, 2.24]$), where our system receives results in the highest tier (*Excellent*), which denotes the range of the best 10% of systems in the benchmark dataset. The categories *Attractiveness* ($M = 1.83$, $\sigma = 0.83$, 95% CI $[1.55, 2.12]$) and *Efficiency* ($M = 1.63$, $\sigma = 0.9$, 95% CI $[1.32, 1.93]$) were ranked in the following tier (*Good*), indicating that 10% of systems in the benchmark dataset scored better while 75% scored worse. *Dependability* ($M = 1.39$, $\sigma = 0.97$, 95% CI $[1.05, 1.72]$) was ranked to be *Above Average*, with 25% of systems in the benchmark dataset

Table 2: Task durations and correctness rates per task type: The first entry inside the brackets is the mean value, the second one the standard deviation, and the third the CI. The minimal and maximal mean values and standard deviations are bold.

	T1: **Attributes**	T2: **Correlations**	T3: **Clusters**	T4: **Outlier**	T5: **No Relation**	**Avrg.**
Completion Time [s]	(**21.77**, **12.06**, [17.66, 25.9])	(50.58, 24.77, [42.13, 59.03])	(37.75, 19.81, [30.99, 44.5])	(**74.34**, **41.12**, [60.31, 88.37])	(29.6, 14.54, [24.65, 34.57])	(42.81, 30.61, [32.36, 53.26])
Correctness [%]	(**99.24**, **4.35**, [0.98, 1.0])	(94.7, 13.63, [0.9, 0.99])	(96.96, 8.28, [0.94, 1.0])	(**87.12**, **17.81**, [0.81, 0.93])	(90.9, 17.48, [0.85, 0.97])	(93.78, 13.91, [0.89, 0.99])

scoring better and 50% scoring worse. The last category *Perspicuity* ($M = 1.106$, $\sigma = 1.09$, 95% CI $[0.73, 1.48]$) got a *Below Average* rating, meaning that 50% of systems in the benchmark dataset scored better, and 25% scored worse.

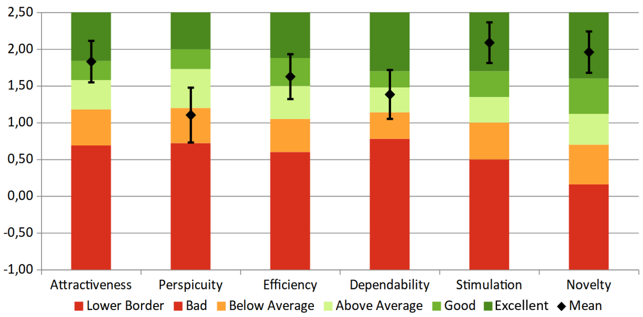

Figure 3: Results of the UEQ in context with the benchmark dataset. The diamonds represent the mean values and the bars the CIs.

To evaluate the system's usability, we took the SUS questionnaire into account [6]. It is composed of ten statements, each having a five-point scale ranging from 1, meaning "strongly disagree", to 5, implying "strongly agree". Since the original work that introduces the SUS does not come along with any interpretation of the score, we take two works into account: Bangor et al. [4] summarize SUS scores by interface type and add an adjective rating scale to the SUS for better interpretation, and Sauro and Lewis [53] introduce a curved grading scale.

The mean SUS score of SPLOCIS ($M = 78.94$, $\sigma = 15.91$, 95% CI [73.51, 84.37]) corresponds to a *Good* ($M = 71.4$, $\sigma = 11.6$) rating and exceeds the overall mean ($M = 69.5$) and the mean of immersive VR interfaces in particular ($M = 72.4$), as reported in the work of Bangor et al. According to the work of Sauro and Lewis, our mean SUS corresponds to the grade "A-" (for which the mean SUS has to be within a range of 78.9 and 80.7), taking additionally the CI into account, we lie between the grades "B-" and "A+".

5.4 Statements Agreement

To get deeper insights into participants' perceptions of our system, they rated custom system-specific statements on a seven-point Likert scale, with 1 indicating "strongly disagree" and 7 "strongly agree", listed in Figure 4. Statements with an odd number (S1, S3, S5, S7) are "positively worded" statements, meaning that ratings closer to 7 indicate a more positive evaluation of our system. Statements having an even number (S2, S4, S6) are "negatively verbalized", meaning the lower the agreement, the better the evaluation. While all mean values and CIs of positive statements are above the general neutral score of 4, the ones for the negative statements are below, indicating an overall positive result.

6 Discussion

Based on the results of our user study, we will answer our three research questions in the following. The FMS scores indicate that our participants were in a good shape to complete the study, suggesting that it is unlikely that cybersickness had a significant influence on the other measurements. Despite two exceptions in which higher scores were reported, the respective participant decided to continue the study after being offered a moment to take a rest. The distributions of prior experiences listed in Table 1 suggest large variations in VR knowledge within the group. While most participants saw their prior data visualization experience at a medium level, no participant considered themselves an expert in the field of scatterplots.

6.1 RQ1 – Influence of Task Type

Our first question was how the task type in the SPLOCIS system affects performance and cognitive load, and if there are task types that are particularly (un)suitable. For this, we take the results of the Raw-TLX, the task correctness, as well as the task completion times (Section 5.2) into account.

As all mean task loads induced by our system were below the reference values of Grier [23] and Hertzum [31], and given the fact that tasks could be solved on average in under 43 s with an average correctness rate of 93.78%, we conclude that the SPLOCIS system allows to solve tasks in a quick and accurate way without imposing unreasonably high task loads. Noticeably, the task type to find plots by their **Attributes** (T1) showed the lowest standard deviations for task completion time and task correctness among the participants, could be solved fastest while leading to the highest correctness rate, and to the second lowest task load among all task categories. This indicates that handling the system mechanics to find plots with specific attributes was perceived as easy and feasible. In contrast, the results for the **Outlier** tasks (T4) spread most and reached the highest values regarding the task load, completion time, and correctness rate, identifying it as the most challenging task type. Even though the data was designed to have clear outliers, as described in Section 4.4, the high amount of data points present in the virtual room might have made it hard to spot one outlying data point within one plot. Another possible explanation is that outlier perception is subjective, resulting in a longer cognitive process. Whereas spotting outliers requires users to inspect single plots on a detailed point-based level, other tasks like finding correlations or clusters require to work on a coarser plot-based level, as those patterns are spread all over a whole plot, making them probably more prominent and therefore easier to detect.

While the positive correlation between task completion time and task load does not imply a causality, it seems apparent that if more time was needed to complete a task, the associated task load in that case tended to be higher, possibly due to a longer or more complex thinking process.

6.2 RQ2 – Usability and Feature Wishes

Our second goal was to investigate how intuitive/accessible/suitable the application is for working with multivariate data by exploring 3D scatterplots and which features would enhance our designed interaction metaphors (or data analysis in VR in general). In order to answer this we take the results of the UEQ, SUS (Section 5.3), and ratings of custom statements (Section 5.4) into account. Additionally, we compare our reported quantitative results to the free-text answers submitted when asking for up to three positive and negative aspects as well as feature wishes.

The *Good* SUS score in combination with the UEQ, in which two categories received *Excellent*, and two *Good* results, give evidence

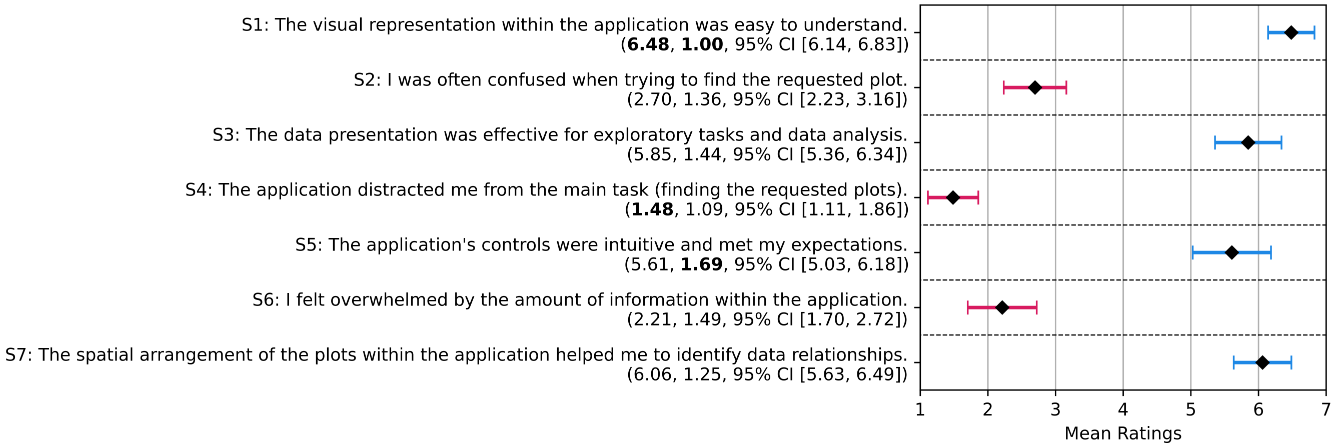

Figure 4: Custom statements agreement: The diamonds represent participants' mean ratings, the bars the CIs. Below each statement the mean value is the first entry listed in the brackets, followed by standard deviation and confidence interval. Minimal and maximal mean and standard deviations are printed bold. 1 means "strongly disagree" and 7 "strongly agree". Odd numbers (S1, S3, S5, S7) are positive statements – a high agreement means a positive result. Even numbers (S2, S4, S6) are negative statements – a low agreement indicates a positive rating.

that the SPLOCIS system has an overall satisfactory usability. This finding is also confirmed by the average grade of "A-" according to the classification scheme of Sauro and Lewis [53].

In particular, the *Excellent* values for *Stimulation* and *Novelty* show a high hedonic quality of our system, implying that it is innovative and exciting to use. Those findings go hand in hand with the free-text answers given when asking about positive aspects: Eleven participants especially highlighted the application's "fun factor", calling it a playful and motivating way to work with data. The high values for *Attractiveness* indicate an overall good impression of the system, also mirrored by the comments in which 23 participants named the whole data visualization as a positive aspect.

The *Good Efficiency* connotes that participants could solve the tasks without unnecessary effort. This is consistent with the positive aspects, in which 29 participants appreciated the system's intuitiveness, understandability, and ease of learning. 14 participants explicitly pointed out the system's efficiency, and eleven liked the clear and well-structured overview the application offered. This is underlined by the high agreement on the data presentation's effectiveness for exploratory tasks and data analysis (see Figure 4, statement S3), and that the spatial arrangement of plots was perceived to facilitate the identification of data relationships (S7). Moreover, most participants stated that the application did not distract them from the main task of finding the requested plot (S4), also underscoring the *Good Efficiency*.

Dependability was *Above Average*, meaning that the users mostly felt in control of the interaction. This is supported by the statement agreement that the application's controls were intuitive and met the participants' expectations (S5), as well as the fact that five emphasized the interactivity as a positive aspect.

With a *Below Average* rating, *Perspicuity* had the lowest values among the UEQ categories, suggesting that it needs some time to get familiar with the system. Despite the fact that 29 stated it to be intuitive and easy to learn, and despite the high agreement that the visual representation was easy to understand (S1), two called the system not intuitive enough when asking for the participants' subjective negative aspects. Most participants stated that they were not often confused when trying to find the requested plots (S2), but it cannot be denied that some confusion was present. Similarly, most participants reported not to feel overwhelmed (S6), but some feelings of being overwhelmed were existent. Additionally, 21 participants saw the physical handling as a possible disadvantage, further specifying that actions happened to be mistaken or accidentally fired. Beyond that, 19 mentioned VR technology in general as a possible drawback, so the needs of having a VR device, getting used to VR technology, and the ever-present probability of suffering from cybersickness, even though our results show that cybersickness was not of significant concern. Moreover, nine mentioned as a negative aspect issues of potentially occluding elements caused by the third spatial dimension and varying perspectives. These widespread results might be explained by the participants' similarly widespread self-reported prior VR and data visualization experience. While some were long-established VR expert users, others had little to no prior experience to either VR or data visualization. We observed that participants new to VR occasionally struggled with memorizing button functionalities. Dealing with an unfamiliar technology in combination with a potentially unfamiliar topic likely contributed to these difficulties, sometimes resulting in confusion or feelings of being overwhelmed. We found a small negative correlation between *Perspicuity* and prior VR experience, $p = -0.29$, which points in the direction of our assumption.

29 participants suggested features eventually enhancing data exploration in 3D immersive space. One mentioned that having access to the user manual at any time would be helpful, which aligns with the comments of accidentally mistaking buttons. Five mentioned different extended rotation possibilities. While two-handed rotations as well as single-handed rotations of grabbed plots were possible (either by rotating physically or by pushing the joystick of the right controller horizontally to induce a rotation of the plot around the y-axis), two mentioned that it would have been helpful to enable a "joystick-induced rotation" around more than just the plot's up-axis in order to not stress the wrist too much. Furthermore, two assumed that rotating all currently highlighted 3D plots within the SPLOC at once would accelerate 3D plot investigation. Four mentioned functionalities to make trends and correlations more directly visible. This ranged from automatically selecting general patterns by the click of a button (e.g., correlations or outliers), or to work with lines or colors to make trends more easily recognizable. Three mentioned "more details on demand"; in particular, one participant suggested that hovering over a 3D plot should display all of its attributes in a small separate UI element. Two would have liked to select individual data points within a plot, with one adding that displaying selected points coordinates would be beneficial, and one suggestion was the possibility to mark plots that seem important.

6.3 RQ3 – Repeated Axis Assignments

To answer the third research question, to what extent plots with repeated axis assignments disturb or support the exploration process, we use the data gathered through the ratings of S6 and S7 (Section 5.4), as well as free-text form answers to custom questions about the use cases participants had in mind for those plots and their reasons to continue showing or excluding them from our system.

First, the quantity of data displayed at a time seems to be accurate as the participants, on average, stated not to feel overwhelmed by the amount of information (S6). In addition, the overall structure seems to be suitable for data exploration tasks due to the high agreement that the spatial arrangement of the plots within the application helped to identify data relationships (S7). Taking this as a basis, we dive deeper into users' opinions regarding the fact that SPLOCIS features plots having the same attributes on multiple axes.

Two third of the participants mentioned possible use cases for these plots, thereby indicating that there is merit in keeping them included in the system. Some mentioned they allow to check whether the individual axis scaling seems appropriate for the range of values within one attribute, or to check if an attribute contains any data, or if someone accidentally uploaded twice the same attribute. Three brought up the aid of orientation, three mentioned that they make comparisons of attributes easier, even if a 3D plot contains the same attribute twice. 13 suggestions were about to keep visualizing them for the sake of completeness and consistency; that empty spots within the SPLOC, resulting from their exclusion, would be irritating. Interestingly, eight participants were of the opposite opinion and considered their existence to be confusing – apparently more than the aforementioned empty spots. This confusion is underlined by the fact that nearly half of all errors were due to the condition violation of choosing plots having distinct attributes, indicating that plots with repeated axis assignments could quickly be grabbed by mistake. While ten considered them to be unnecessary, as the information within those plots can be obtained somewhere else instead, eight valued them for gathering an attribute's data distribution in isolation, as they reveal the variance, outliers, data scales, potential clustering of points, and provide a complete picture of the data distribution. Seven participants suggested a toggle mechanism that allows to visualize those plots if needed and exclude them otherwise which would probably satisfy both of these distinct opinions.

6.4 Limitations

In the following we like to mention some limitations of our user study to motivate relevant aspects of future work. First of all, we note that the gender distribution among our participants was not balanced. On top of that, there were substantial variations among the participants regarding prior VR experience, and they were not professional data analysts. Furthermore, we worked with artificial datasets to have full control over the patterns present in the data – that raises question if and how real-world datasets would affect the results. In addition, we defined specific thresholds to distinguish between valid and invalid plots when asking for patterns. Here, it is important to note that those thresholds are not entirely objective. E.g., some users might have classified more subtle deviating data points as outliers, which we did not categorize to be ones. Lastly, our study involved a single condition and relied on a theoretical comparison with ImAxes rather than an empirical one. Even though our user study was mainly a proof of concept, all of these aspects should be considered when interpreting the findings.

7 Conclusion and Future Work

This work did the first step in answering the higher-level research question of how to make 3D scatterplots effectively, comfortably and intuitively explorable in iVEs. Based on best practices from HCI we developed SPLOCIS, an immersive system designed for investigating multivariate data through 2D and 3D scatterplot exploration. Collectively, despite the participants' wide range of knowledge levels regarding VR technology and scatterplots as a data visualization method, all participants were able to use the SPLOCIS system to identify plots by their **Attributes** (T1), find plots having **Correlations** (T2), **Clusters** (T3), an **Outlier** (T4) or **No Relation** (T5) in an overall quick, accurate, and comfortable way. Not all tasks were equally classified in their difficulty, time investment and error-proneness; while finding plots by **Attributes** (T1) appeared to be the easiest, detecting plots having **Outliers** (T4) seemed to be the most challenging task. Moreover, SPLOCIS was mostly perceived to be intuitive, accessible and well-suited for working with multivariate data in a novel, stimulating, efficient, and attractive way. Some challenges likely resulted from the still present novelty and skepticism of VR technology. At the same time, the suggested new features offer valuable opportunities for making immersive data exploration even more suitable and comfortable. Plots with repeated axis assignments were perceived controversial: While some found them to be useful, some called them confusing. A solution satisfying both views is to add a customization option, enabling users to show or hide them when needed.

Future work should focus on integrating advanced interactions like brushing and linking or dynamic filtering besides participants' suggested new features, followed by an evaluation of an updated SPLOCIS version with domain experts, as our group of participants had different backgrounds and varying experience levels, as explained in Section 6.4. Furthermore, it makes sense to perform comparison studies with different immersive techniques like ImAxes [13] or other 2D (non-immersive) tools to find out the individual strengths and weaknesses of the respective systems.

While the rationale behind SPLOCIS was to save operational steps to build a data visualization offering an overview, some participants felt overwhelmed by this high amount of data. In contrast to that, ImAxes very likely avoids this feeling as users build their own data visualizations "from scratch", which takes some time and operational steps, but also implies that users are not confronted with an already existent large data visualization. Additionally, as SPLOCIS is based on IATK [12], it inherits the same limitations regarding the number of data points that can be displayed. On top of that, increasing the number of attributes enlarges the SPLOMs, potentially resulting in excessively large scenes. While reducing the plots' sizes is possible, patterns and labels must remain legible, resulting in a trade-off. Besides that, increasing the number of attributes and data points can cause/facilitate visual clutter and overplotting. Therefore, it makes sense to evaluate the applicability of recent approaches mentioned in Section 2.1 that aim to enhance the interpretability and scalability of scatterplots and SPLOMs. Overall, we believe that immersive VR offers several novel opportunities to explore and understand complex datasets interactively, the potential of which is still to be exploited more comprehensively.

Acknowledgments

This work was supported by "Stiftung Innovation in der Hochschullehre" through the project "Hybrid Learning Center" (FBM2020-EA-690-01130). We appreciate the assistance of Mikel D. Jedrusiak and Gerrit Nolte in elaborating the interaction formalization.

References

[1] Meta Quest 3. https://www.meta.com/de/quest/quest-3/. Accessed: 2025-08-29. 3

[2] XR Interaction Toolkit. https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@3.2. Accessed: 2025-08-29. 4

[3] D. Allcoat and A. von Mühlenen. Learning in virtual reality: Effects on performance, emotion and engagement. *Research in Learning Technology*, 26, 2018. doi: 10.25304/rlt.v26.2140 3

[4] A. Bangor, P. Kortum, and J. Miller. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of usability studies*, 4(3):114–123, 2009. 7

[5] A. Batch, A. Cunningham, M. Cordeil, N. Elmqvist, T. Dwyer, B. H. Thomas, and K. Marriott. There Is No Spoon: Evaluating Performance, Space Use, and Presence with Expert Domain Users in Immersive Analytics. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):536–546, 2020. doi: 10.1109/TVCG.2019.2934803 3

[6] J. Brooke. SUS: A 'Quick and Dirty' Usability Scale. In *Usability Evaluation In Industry*, chap. 21, pp. 189–194. CRC Press, London, 1 ed., 1996. doi: 10.1201/9781498710411 5, 7

[7] D. B. Carr, R. J. Littlefield, W. L. Nicholson, and J. S. Littlefield. Scatterplot Matrix Techniques for Large N. *Journal of the American Statistical Association*, 82(398):424–436, 1987. doi: 10.1080/01621459.1987.10478445 2

[8] J. M. Chambers, W. S. Cleveland, and P. A. Tukey. *Graphical Methods for Data Analysis (1st ed.)*. CRC Press, 1983. doi: 10.1201/9781351072304 1, 2

[9] C. Chang, T. Dwyer, and K. Marriott. An Evaluation of Perceptually Complementary Views for Multivariate Data. In *2018 IEEE Pacific Visualization Symposium*, pp. 195–204, 2018. doi: 10.1109/PacificVis.2018.00033 2

[10] M. R. Chester, M. J. Rys, and S. A. Konz. Leg swelling, comfort and fatigue when sitting, standing, and sit/standing. *International Journal of Industrial Ergonomics*, 29(5):289–296, 2002. doi: 10.1016/S0169-8141(01)00069-5 3

[11] J. Cohen. *Statistical power analysis for the behavioral sciences (2nd ed.* L. Erlbaum Associates, 1988. 6

[12] M. Cordeil, A. Cunningham, B. Bach, C. Hurter, B. H. Thomas, K. Marriott, and T. Dwyer. IATK: An Immersive Analytics Toolkit. In *IEEE Conference on Virtual Reality and 3D User Interfaces*, pp. 200–209, 2019. doi: 10.1109/VR.2019.8797978 3, 9

[13] M. Cordeil, A. Cunningham, T. Dwyer, B. H. Thomas, and K. Marriott. ImAxes: Immersive Axes as Embodied Affordances for Interactive Multivariate Data Visualisation. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, pp. 71–83. Association for Computing Machinery, 2017. doi: 10.1145/3126594.3126613 3, 9

[14] T. N. Dang and L. Wilkinson. ScagExplorer: Exploring Scatterplots by Their Scagnostics. In *IEEE Pacific Visualization Symposium*, pp. 73–80, 2014. doi: 10.1109/PacificVis.2014.42 2

[15] R. De Maesschalck, D. Jouan-Rimbaud, and D. Massart. The Mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems*, 50(1):1–18, 2000. doi: 10.1016/S0169-7439(99)00047-7 5

[16] B. Dy, N. Ibrahim, A. Poorthuis, and S. Joyce. Improving visualization design for effective multi-objective decision making. *IEEE Transactions on Visualization and Computer Graphics*, 28(10):3405–3416, 2022. doi: 10.1109/TVCG.2021.3065126 2

[17] U. Engelke, J. Vuong, and J. Heinrich. Visual Performance in Multidimensional Data Characterisation with Scatterplots and Parallel Coordinates. *Electronic Imaging*, 28(16):1–1, 2016. doi: 10.2352/ISSN.2470-1173.2016.16.HVEI-136 2

[18] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining*, vol. 96, pp. 226–231. AAAI Press, 1996. 5

[19] R. Etemadpour, E. Monson, and L. Linsen. The Effect of Stereoscopic Immersive Environments on Projection-Based Multi-dimensional Data Visualization. In *International Conference on Information Visualisation*, pp. 389–397, 2013. doi: 10.1109/IV.2013.51 1, 3

[20] Y. Farmani and R. J. Teather. Evaluating discrete viewpoint control to reduce cybersickness in virtual reality. *Virtual Reality*, 24(4):645–664, 2020. 4

[21] H. Garner and S. J. Fernstad. Comparative evaluation of the Scatter Plot Matrix and Parallel Coordinates Plot Matrix. In *26th International Conference Information Visualisation*, pp. 114–122, 2022. doi: 10.1109/IV56949.2022.00027 1

[22] A. Gracia, S. González, 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 of quality quantification. *Information Visualization*, 15(1):3–30, 2016. doi: 10.1177/1473871614556393 2

[23] R. A. Grier. How high is high? A meta-analysis of NASA-TLX global workload scores. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 59, pp. 1727–1731, 2015. 6, 7

[24] L. Harrison, F. Yang, S. Franconeri, and R. Chang. Ranking Visualizations of Correlation Using Weber's Law. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1943–1952, 2014. doi: 10.1109/TVCG.2014.2346979 1, 2

[25] S. G. Hart. NASA-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomic Society annual meeting*, 50(9):904–908, 2006. doi: 10.1177/154193120605000909 5

[26] S. G. Hart and L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, vol. 52 of *Advances in Psychology*, pp. 139–183. North-Holland, 1988. doi: 10.1016/S0166-4115(08)62386-9 5, 6

[27] D. Hayatpur, H. Xia, and D. Wigdor. DataHop: Spatial Data Exploration in Virtual Reality. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, UIST '20, pp. 818–828. Association for Computing Machinery, 2020. doi: 10.1145/3379337.3415878 3

[28] N. Heckert, J. Filliben, C. Croarkin, B. Hembree, W. Guthrie, P. Tobias, and J. Prinz. Handbook 151: NIST/SEMATECH e-Handbook of Statistical Methods, 2002. doi: 10.18434/M32189 1, 5

[29] M. Henley, M. Hagen, and R. D. Bergeron. Evaluating Two Visualization Techniques for Genome Comparison. In *11th International Conference Information Visualization*, pp. 551–558, 2007. doi: 10.1109/IV.2007.47 2

[30] N. Henze and B. Zirkler. A class of invariant consistent tests for multivariate normality. *Communications in Statistics - Theory and Methods*, 19(10):3595–3617, 1990. doi: 10.1080/03610929008830400 5

[31] M. Hertzum. Reference values and subscale patterns for the task load index (TLX): a meta-analytic review. *Ergonomics*, 64(7):869–878, 2021. doi: 10.1080/00140139.2021.1876927 6, 7

[32] C. Hurter, N. H. Riche, S. M. Drucker, M. Cordeil, R. Alligier, and R. Vuillemot. FiberClay: Sculpting Three Dimensional Trajectories to Reveal Structural Insights. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):704–714, 2019. doi: 10.1109/TVCG.2018.2865191 3

[33] R. Kanjanabose, A. Abdul-Rahman, and M. Chen. A Multi-task Comparative Study on Scatter Plots and Parallel Coordinates Plots. *Computer Graphics Forum*, 34(3):261–270, 2015. doi: 10.1111/cgf.12638 2

[34] B. Keshavarz and H. Hecht. Validating an Efficient Method to Quantify Motion Sickness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 53(4):415–426, 2011. doi: 10.1177/0018720811403736 5

[35] M. Kraus, N. Weiler, D. Oelke, J. Kehrer, D. A. Keim, and J. Fuchs. The Impact of Immersion on Cluster Identification Tasks. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):525–535, 2020. doi: 10.1109/TVCG.2019.2934395 1, 3

[36] X. Kuang, H. Zhang, S. Zhao, and M. McGuffin. Tracing Tuples Across Dimensions: A Comparison of Scatterplots and Parallel Coordinate Plots. *Computer Graphics Forum*, 31(3pt4):1365–1374, 2012. doi: 10.1111/j.1467-8659.2012.03129.x 2

[37] B. Laugwitz, T. Held, and M. Schrepp. Construction and Evaluation of a User Experience Questionnaire. In A. Holzinger, ed., *HCI and Usability for Education and Work*, pp. 63–76. Springer Berlin Heidelberg, 2008. doi: 10.1007/978-3-540-89350-9_6 5

[38] B. Lee, M. Cordeil, A. Prouzeau, and T. Dwyer. FIESTA: A Free Roaming Collaborative Immersive Analytics System. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*, pp. 335–338. Association for Computing Machinery, 2019. doi: 10.1145/3343055.3360746 3

[39] L. Lisle, K. Davidson, E. J. Gitre, C. North, and D. A. Bowman.

Sensemaking Strategies with Immersive Space to Think. In *IEEE Virtual Reality and 3D User Interfaces*, pp. 529–537, 2021. doi: 10.1109/VR50410.2021.00077 4

[40] L. Lisle, K. Davidson, L. Pavanatto, I. A. Tahmid, C. North, and D. A. Bowman. Spaces to Think: A Comparison of Small, Large, and Immersive Displays for the Sensemaking Process. In *IEEE International Symposium on Mixed and Augmented Reality*, pp. 1084–1093, 2023. doi: 10.1109/ISMAR59233.2023.00125 4

[41] 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 1

[42] A. Mayorga and M. Gleicher. Splatterplots: Overcoming Overdraw in Scatter Plots. *IEEE Transactions on Visualization and Computer Graphics*, 19(9):1526–1538, 2013. doi: 10.1109/TVCG.2013.65 2

[43] J. P. McIntire and K. K. Liggett. The (Possible) Utility of Stereoscopic 3D Displays for Information Visualization: The Good, the Bad, and the Ugly. In *IEEE VIS International Workshop on 3DVis*, pp. 1–9, 2014. doi: 10.1109/3DVis.2014.7160093 3

[44] O. Merhi, E. Faugloire, M. Flanagan, and T. A. Stoffregen. Motion Sickness, Console Video Games, and Head-Mounted Displays. *Human Factors*, 49(5):920–934, 2007. doi: 10.1518/001872007X230262 3

[45] T. Munzner. *Visualization Analysis and Design*. CRC press, 2014. 2, 4

[46] R. Netzel, J. Vuong, U. Engelke, S. O'Donoghue, D. Weiskopf, and J. Heinrich. Comparative eye-tracking evaluation of scatterplots and parallel coordinates. *Visual Informatics*, 1(2):118–131, 2017. doi: 10.1016/j.visinf.2017.11.001 2

[47] W. Peng, M. Ward, and E. Rundensteiner. Clutter Reduction in Multi-Dimensional Data Visualization Using Dimension Reordering. In *IEEE Symposium on Information Visualization*, pp. 89–96, 2004. doi: 10.1109/INFVIS.2004.15 4

[48] M. D. Plumlee and C. Ware. Zooming versus multiple window interfaces: Cognitive costs of visual comparisons. *ACM Transactions on Computer-Human Interaction*, 13(2):179–209, 2006. doi: 10.1145/1165734.1165736 4

[49] J. Poco, R. Etemadpour, F. Paulovich, T. Long, P. Rosenthal, M. Oliveira, L. Linsen, and R. Minghim. A Framework for Exploring Multidimensional Data with 3D Projections. *Computer Graphics Forum*, 30(3):1111–1120, 2011. doi: 10.1111/j.1467-8659.2011.01960.x 3

[50] G. E. Riccio and T. A. Stoffregen. An Ecological Theory of Motion Sickness and Postural Instability. *Ecological Psychology*, 3(3):195–240, 1991. doi: 10.1207/s15326969eco0303_2 3

[51] A. Sarikaya and M. Gleicher. Scatterplots: Tasks, Data, and Designs. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):402–412, 2018. doi: 10.1109/TVCG.2017.2744184 2, 5

[52] K. A. Satriadi, B. Ens, M. Cordeil, T. Czauderna, and B. Jenny. Maps Around Me: 3D Multiview Layouts in Immersive Spaces. *Proc. ACM Hum.-Comput. Interact.*, 4, 2020. doi: 10.1145/3427329 4

[53] J. Sauro and J. R. Lewis. *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann, 2nd ed., 2016. 7, 8

[54] B. Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In B. B. Bederson and B. Shneiderman, eds., *The Craft of Information Visualization*, Interactive Technologies, pp. 364–371. Morgan Kaufmann, 2003. doi: 10.1016/B978-155860915-0/50046-9 2, 4

[55] A. Tatu, P. Bak, E. Bertini, D. Keim, and J. Schneidewind. Visual quality metrics and human perception: an initial study on 2D projections of large multidimensional data. In *Proceedings of the International Conference on Advanced Visual Interfaces*, pp. 49–56. Association for Computing Machinery, 2010. doi: 10.1145/1842993.1843002 2

[56] E. R. Tufte. Envisioning information. *Optometry and Vision Science*, 68(4):322–324, 1991. 2, 4

[57] S. van den Elzen and J. J. van Wijk. Small Multiples, Large Singles: A New Approach for Visual Data Exploration. *Computer Graphics Forum*, 32:191–200, 2013. doi: 10.1111/cgf.12106 4

[58] J. A. Wagner Filho, C. Freitas, and L. Nedel. VirtualDesk: A Comfortable and Efficient Immersive Information Visualization Approach. *Computer Graphics Forum*, 37(3):415–426, 2018. doi: 10.1111/cgf.13430 3

[59] J. A. Wagner Filho, M. F. Rey, C. M. D. S. Freitas, and L. Nedel. Immersive Visualization of Abstract Information: An Evaluation on Dimensionally-Reduced Data Scatterplots. In *IEEE Conference on Virtual Reality and 3D User Interfaces*, pp. 483–490, 2018. doi: 10.1109/VR.2018.8447558 1, 3

[60] L. Wilkinson, A. Anand, and R. Grossman. Graph-Theoretic Scagnostics. In *IEEE Symposium on Information Visualization*, pp. 157–164, 2005. doi: 10.1109/INFVIS.2005.1532142 2

[61] L. Wilkinson, A. Anand, and R. Grossman. High-dimensional visual analytics: Interactive exploration guided by pairwise views of point distributions. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1363–1372, 2006. doi: 10.1109/TVCG.2006.94 2

[62] J. Yang, A. Patro, S. Huang, N. Mehta, M. Ward, and E. Rundensteiner. Value and Relation Display for Interactive Exploration of High Dimensional Datasets. In *IEEE Symposium on Information Visualization*, pp. 73–80, 2004. doi: 10.1109/INFVIS.2004.71 2

[63] A. Yates, A. Webb, M. Sharpnack, H. Chamberlin, K. Huang, and R. Machiraju. Visualizing Multidimensional Data with Glyph SPLOMs. *Computer Graphics Forum*, 33(3):301–310, 2014. doi: 10.1111/cgf.12386 2