There Is No Spoon: Evaluating Performance, Space Use, and Presence with Expert Domain Users in Immersive Analytics

Andrea Batch, Andrew Cunningham, Maxime Cordeil, Niklas Elmqvist, *Senior Member, IEEE*, Tim Dwyer, Bruce H. Thomas, *Senior Member, IEEE*, Kim Marriott

Fig. 1: Macroeconomics analysis in the ImAxes immersive analytics tool [11]. (Photo by Samuel Zeller on Unsplash.)

Abstract—Immersive analytics turns the very space surrounding the user into a canvas for data analysis, supporting human cognitive abilities in myriad ways. We present the results of a design study, contextual inquiry, and longitudinal evaluation involving professional economists using a Virtual Reality (VR) system for multidimensional visualization to explore actual economic data. Results from our preregistered evaluation highlight the varied use of space depending on context (exploration vs. presentation), the organization of space to support work, and the impact of immersion on navigation and orientation in the 3D analysis space.

Index Terms—Design study, evaluation, economic analysis, immersive analytics.

✦

1 INTRODUCTION

THE ECONOMIST pivots on the balls of her feet and reaches out, her heart thumping like a war drum in her chest as she deftly grabs the missing data dimension,[1] a precious jewel glimmering in the austere analytics environment. "Gotcha," she murmurs to herself. She's almost solved it now! She wipes the sweat trickling down her brow from under the headset with the back of her other hand while she idly twirls the glowing 3D axis she's holding and scans the galaxy of data displays arrayed in front of her. Now where is that net exports scatterplot... there! With a sense of satisfaction, she slides the erring axis into place next to the plot, instantly turning it into a scatterplot matrix.[2] Then she steps back and admires her handiwork.[3] All has been revealed.

> **Actions:** [1]axis control; [2]scatterplot matrix; [3]physical navigation.

- *Andrea Batch and Niklas Elmqvist are with the University of Maryland in College Park, MD, USA. E-mail: {ajulca, elm}@umd.edu.*
- *Andrew Cunningham and Bruce H. Thomas are with the University of South Australia in Adelaide, Australia. E-mail: {andrew.cunningham, bruce.thomas}@unisa.edu.au.*
- *Maxime Cordeil, Tim Dwyer, and Kim Marriott are with Monash University in Melbourne, Australia. E-mail: {tim.dwyer, kim.marriott}@monash.edu.*

☙ ❧

In this paper, we report on a design study on the use of Immersive Analytics (IA) [28] in Virtual Reality (VR) for professional economic analysis in a U.S. federal agency. Inspired by Sedlmair's design study methodology [38], this overall study consisted of multiple phases:

1. A *design stage* where we collected requirements using contextual inquiry methodology [6] and improved an existing immersive VR system for multidimensional data analysis—ImAxes [11]—to support macroeconomics data;

2. A *formative "in-the-wild" deployment* of the prototype application in a communal space, which lead to multiple incremental insights and improvements of the prototype; and

3. An *in-depth mixed methods study* (preregistered) involving professional economic analysts exploring their own datasets in our immersive economics environment, and then presenting their findings to the experiment administrator.

The results from these studies include observations, video and audio recordings, interaction logs, and subjective interview plus survey feedback from the participants. In particular, we report on the use and organization of space to support analysis and presentation, barriers against effective use of immersive environments for data analysis, and the impact of immersion on navigation and orientation in 3D.

We target the macroeconomics use case specifically because its analyst population is typical of the professional analysts that many immersive analytics applications purport to support. Economic analysis is characterized by large-scale, high-dimensional, and abstract data—often with a temporal component—that typically is visualized in many separate views [5], which makes it particularly amenable to an immersive setting. That is, we anticipate a future where headset devices make

screens obsolete and provide data analysts—such as economists—the ability to arrange their work in the spaces around them in a way that is impossible with desktop interaction. For now we use VR headsets, which currently offer the best immersive visualization capabilities (field of view and resolution), but we expect that the findings in our study about use of space will be transferable to future devices, such as augmented reality, as discussed in Section 7.

In general, our results are consistent with prior art on the organization of space to simplify choice, perception, and computation [22], as well as the spatial arrangement of digital artifacts to support recall, cognition, and clustering [1]. We found that during initial data exploration our participants would primarily arrange views in an egocentric fashion around themselves during personal exploration, opportunistically placing new views in the closest free space. The arrangement was more considered when participants presented their findings. Strategies included careful curation of the exploration views and arranging views into a chronological narrative for presentation to a third party. Furthermore, presentations would often involve more complex visualizations, whereas the exploratory phase was characterized by the creation of many transient visualizations that were quickly discarded. We were surprised to find that many participants created more complex non-traditional visualizations such as 3D scatterplots. We were also surprised that well-known limitations of VR such as fatigue or text legibility were not of significant concern. Participants were overwhelmingly positive about their experience, even those unfamiliar with VR. They reported a higher level of engagement and presence in the VR environment than in a traditional desktop environment, and they also found that creating visualizations was faster and easier.

We claim the following contributions: (*i*) the first mixed methods study of the concept of embodied immersive analytics (Sec. 3), as described in the original ImAxes paper [11]; (*ii*) improvements to ImAxes making it suitable for macroeconomics, including several distinct new features designed for this particular application (Sec. 4.4); (*iii*) formative feedback on a deployment of ImAxes "in the wild" at a U.S. federal agency (Sec. 4); and (*iv*) performance, behaviorial, and subjective results from an in-depth evaluation with professional economic analysts from the same agency (Sec. 5 and 6).

❧ ☙

CURIOSITY piqued about the relationship between consumer expenditure on recreational goods and spending on food services, the economist grabs the two variables off a rotating shelf,[4] sticks them together,[5] and immediately sees the correlation between the two. But how does it vary over time? She plucks a temporal axis that she'd previously tucked away, adds it to the plot with a flick of the wrist,[6] and spins it[7] to get a quick impression of how the dynamic between the commodities has changed over the last thirty years. There is no one to interrupt her; no nosy coworker poking his head over her cubicle wall to talk about the latest sporting event, the animated conversation between colleagues up the corridor invisible and all but inaudible to her. Just her alone with her data.[8] Bliss.

Actions: [4]Lazy Susan axis selector; [5]2D scatterplot; [6]3D scatterplot; [7]visualization manipulation; [8]immersive economics.

---

2 RELATED WORK

Virtual (VR), Mixed (MR), and Augmented Reality (AR)—immersive display and input technologies on the reality-virtuality continuum [29]—have long been used for visualizing physically embedded data [35, 24, 25, 49]. Recently, this has been extended to include more abstract data using *immersive analytics* [9, 15, 28].

2.1 Immersive Analytics

According to Dwyer et al. [15], *"Immersive Analytics is the use of engaging, embodied analysis tools to support data understanding and*

decision making." IA applications tend to be based on immersive output and input technologies—e.g., VR, AR, or MR—for the purpose of evoking engaging and embodied analysis experiences. The implication is that these immersive technologies can now be effectively used for any form of data, including abstract, non-spatial data, i.e., what Bowman *et al.* call Information-Rich Virtual Environments [7].

Several IA applications are emerging that leverage the presence and engagement of VR. Simpson *et al.* [44] proposed an IA tool to explore climate economy models by leveraging spatial understanding from immersion on 2D multidimensional representations. The open-source ImAxes system [11], which we extend in this work, presents the concept of an embodied axis to enable users to quickly build multidimensional visualizations in VR using natural interactions. FiberClay [17] uses an immersive approach for exploring large-scale spatial trajectory data in 3D, and the system was informally evaluated with air traffic controllers. However, none of these systems involved formal studies on how experts use the available 3D space, or how they might use immersive systems in day-to-day data analysis.

Butscher *et al.* [8] proposed the ART tool for collaborative AR parallel-coordinate-plot viewing with tabletop touch-input. They performed an informal group-based walkthrough evaluation of the system with expert users exploring immersion, presence, spatial layout, and engagement, whereas our study involved individual participants in hands-on work with their own data. Thus, their findings—while valuable and formative for our work—focused more on collaboration.

In summary, the premise of IA is that the immersive setting will yield a richer and more embodied data analysis experience than traditional means. IA has been touted to decrease level of indirection, allow more natural input mechanisms, and the free-form space of a 3D virtual environment, which enables intelligent space usage [1, 22]. However, we are aware of no empirical studies that test these factors for IA, and thus we conducted the present study.

2.2 On the Use of Space in 2D and 3D

Managing and navigating space, virtual or physical alike, has always been central to human cognition. As Norman holds, *"it is things that make us smart"* [32], and according to distributed [18], embodied [39], and extended [10] forms of cognition, this very much includes physical space. In seminal work from cognitive psychology, Kirsh [22] demonstrated that humans tend to offload cognitive tasks in physical space to simplify choice, perception, and internal computation. But how many of these ideas translate to digital space on a computer screen?

Kirsh and Maglio [23] showed that screen space can support internal computation in so-called *epistemic* actions—actions that serve no other purpose than to facilitate thought—in the video game Tetris. Similar effects have also been observed for recall through *spatial memory*: using the Data Mountain [36], where digital objects are arranged on the face of a pseudo-3D "mountain," participants were able to find previously placed website icons significantly faster than when using a conventional bookmark display. This harnessing of spatial memory is also similar to users leveraging *physical navigation* [3] in large and immersive displays with persistent locations of objects, thus allowing muscle memory and proprioception to replace some of the mental effort involved in spatial navigation.

In particular, having access to large visual spaces has been shown to be useful for analytical tasks. For example, screen space can be organized into complex structures such as lists, stacks, heaps, and composites [42], thus reducing the need for mental models. Tan et al. [47] compared analytical task performance between monitors and wall displays, and showed that a physically large display yields significant improvements due to the increased immersion and presence, which biased participants to adopt an egocentric view of the data. Reda et al. [34] built on such findings to study the impact of physical size on actual visual exploration of data, and found consistent effects where more pixels yielded more discoveries and insights. Finally, Andrews et al. [1] directly addressed strategies for spatial arrangement of documents on a large, tiled 2D display in a visual analytics task. Their observations unearthed several interesting phenomena, such as the support of external memory, the structuring of the space using grouping

and layout, and the high degree of integration between process, representation, and data that the large display space scaffolded.

2.3 Presence, Immersion, and Embodiment

Presence is the subjective psychological experience of being in a virtual or remote space, and *immersion* is the objective characteristics of the technology used to present the space [21]. The sense of *embodiment* refers to the sensations that accrue while being inside, having, and controlling a body in VR. A common method of measuring presence is with questionnaires [4, 26, 37, 45, 50, 51]. These studies make a distinction between immersion and presence, where immersion is a necessary (but not sufficient) condition for the experience of presence in a VR interface [16, 45, 51]. Another necessary condition is *involvement* (or attention): the internal processes and external conditions influencing the user's ability to focus on stimuli in the environment [51]. Clearly, while immersion is tied to the technology used to deliver the virtual environment, presence is a more holistic property that is harder to pin down. Witmer and Singer argue, backed by other foundational research on the subject, that immersion and presence are determined by factors influencing the user's sense of control, realism, sensory feedback/stimulation, and distraction [26, 40, 51].

2.4 Immersive Evaluation "in the Wild"

Our work deploys VR for data analysis in a field setting, where measuring performance in addition to presence and immersion is of special interest. Only a few recent studies exist that study VR "in the wild," and even fewer exist for multidimensional data visualization. For example, Steed et al. [46] performed a study over the web with Samsung Gear VR and Google Cardboard, and they found tangible evidence of aspects of presence and immersion being measurable in this setting. Mottelson and Hornbæk [30] conducted a similar field-deployed evaluation with cardboard VR devices, comparing the results to a laboratory study. Their findings are consistent with those of Steed et al., yet also indicate that performance is impacted by the quality of the VR technology and the internal validity of the study.

While our work was deployed in the field, we retained an embedded researcher, making our work closer to a MILCS (Multiple In-depth Case Studies) study [43]. In contrast, most of the above "in the wild" studies do rely on truly uncontrolled environments. Nevertheless, many of their findings are formative for our study design.

2.5 Cooperative and Contextual Inquiry for Visualization

Cooperative inquiry is a qualitative evaluation method based on an iterative cycle of three primary steps: contextual inquiry, participatory design, and technology immersion [13]. Contextual inquiry is the data collection process in which the researcher and participant form a partnership to reach a shared understanding of the user's experience as part of a broader design study [6, 52]. A recent study [5] employed contextual inquiry to understand data scientist workflows and their relationship to interactive visualization through in-depth interview sessions.

In participatory design, the user partners with the researcher to continuously develop new prototypes for the implementation. One method that we particularly draw from participatory design is to embed a researcher with both the users and designers of the system to act as a *values lever*: A link between user and researcher team who is responsible for translating user requests into technical specifications [41]. On an operational level, this is similar to the *pair analytics* approach proposed by Arias-Hernandez et al. [2], where a visual analytics expert "drives" the system while a domain expert gives directions.

3 STUDY METHODS

Our study involved four main phases (Figure 2): a pilot study (P), a formative "in-the-wild" phase (F), and two in-depth phases (S1+S2).

3.1 Setting and Participant Pool

All phases of the study were conducted at a U.S. federal agency where one of the authors was embedded. The participant pool for all experiments thus consisted of data scientists, economic analysts, and
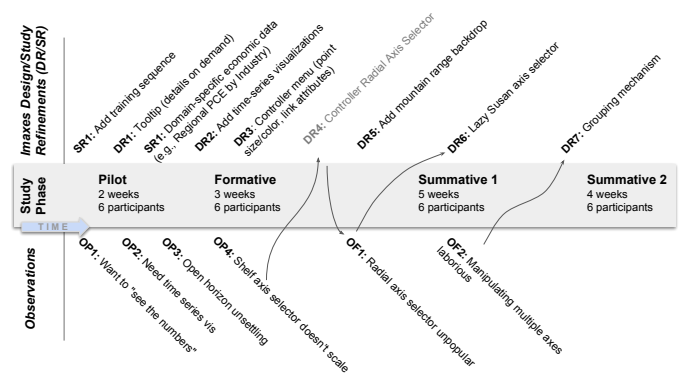

Fig. 2: Process, timeline, refinements, and observations for our study. DR4 shown in grey, as it was replaced based on user feedback.

economists employed, interning, or contracting at this agency. Overall, the education level was high among our participant pool, with all participants having advanced degrees in economics (collectively, 6 master's degree and 3 Ph.D.s), statistics/mathematics (1 master's, 2 Ph.D.s), public policy (2 master's), political science (1 master's, 1 Ph.D.) or similar domains. Participants in our individual in-depth experiments were screened to be experts in data analysis; they routinely used data management and analysis operations daily and had several years of experience working in this duty.

3.2 Apparatus

All studies were conducted in a small office of approximately 10×10 feet (3×3 meters) dedicated to this study. The computing equipment was a personal computer equipped with a Nvidia GeForce GTX 1060 (6GB) GPU, Intel Xeon E5-2620 v3 (2.40GHz) CPU, and 16GB RAM, and running Microsoft Windows 10. The rig was equipped with an HTC Vive VR system, including a head-mounted display (HMD), two base stations, and a monitor that enabled the experimenter to observe the viewpoint of the HMD. The ImAxes application was built using Unity 5.6.5f1. Additional evaluation of video and telemetry data was conducted using a PC equipped with an EVGA GeForce GTX 1080 SC (8 GB) GPU, Intel Core i7-7700 CPU (3.60GHz, 4 cores), and 24 GB RAM, also running Windows 10.

3.3 Data Collection

Here we review the data collection methods employed across all studies. We use the identifiers P (pilot), F (formative), S1, and S2 (summative 1 and 2) to match collection methods to specific phases:

▷ **Demographics Survey (P, S1, S2):** We began our sessions by introducing the study and gathering demographics information. Specifically, we used a written survey to inquire about their past use of VR, their past use of visualizations, gaming experience, and their professional and academic experience.

▷ **Telemetry Recordings (P, F, S1, S2):** The software was instrumented to record controller and headset tracking data over time. The system also recorded specific interactions, such as grabbing and manipulating axes, creating visualizations, and selecting data.

▷ **Video Recordings (P, F, S1, S2):** Two Raspberry Pi Zeros with 8MP Pi cameras and with MotionEyeOS served as webcams set up to capture the interaction space whenever the software was active. One Raspberry Pi was positioned at chest height directly in front of the user's starting position, and the second was positioned in a top corner of the room, a location it shared with one of the Vive's base stations.

▷ **Screen Recordings (S2):** Screen activities were captured using the Windows built-in screen recorder from the game bar. These showed the virtual environment from the participant's viewpoint.

▷ **Audio Recordings (S1, S2):** We recorded participant think-aloud utterances during in-depth sessions using a mobile device.

▷ **Exit Interview (S1, S2):** We ended sessions with a survey and an open-ended interview; answers were recorded and transcribed.

3.4 Common Procedure

Users signed in to use the device. Users were only permitted to access the system if they were formal participants of the study who had signed the consent form. Participants were verbally informed that their activities would be recorded even if the researcher was not present during their use of the implementation. At the end of the study, participants were asked to complete an exit interview and a survey. The procedure for S1 and S2 in particular is given in detail in the study preregistration: `https://osf.io/53e7n/`

3.5 Data Analysis

Our collected data was analyzed with several common methods across the different phases. Here we describe these methods in detail.

3.5.1 Visualization of Spatial Activity

The tracked 3D telemetry data over time provides important insights in how participants move around (physical navigation), interact with 3D objects (axes and visualizations), and arrange their space. To best analyze and present this data, we aggregate movement data over time into a projected 2D grid of the space. We use a top-down view to study physical navigation as well as spatial arrangements of views and axes (heatmap), and a side view to explore interaction heights (histogram).

3.5.2 Replaying Participant Sessions

By combining telemetry data and interaction logs, we are able to replay individual participant sessions. This allows us to understand the participant's view of the analytical space at any point in time. This ability to replay sessions is useful for understanding dynamic behavior and to recreate the arrangement of the space at different times.
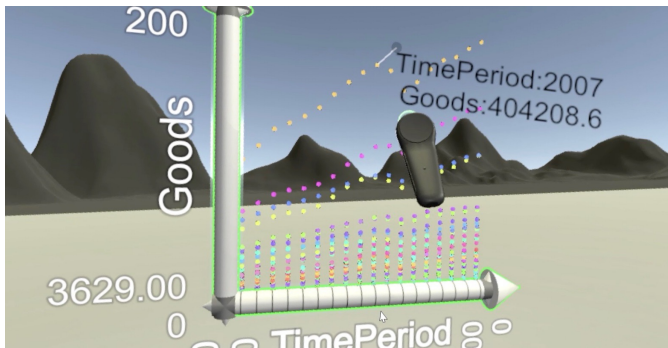

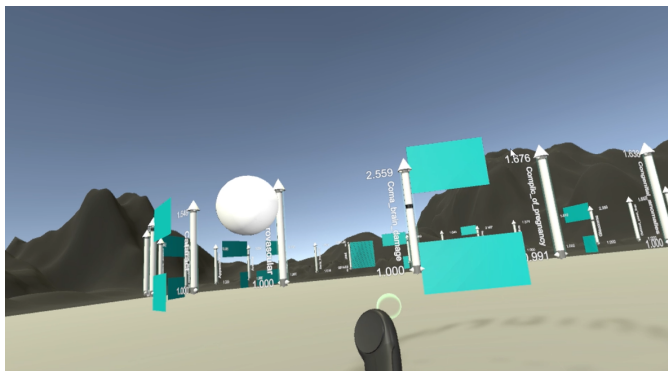Fig. 3: Tooltip providing details-on-demand for data items.

Fig. 4: Lazy Susan menu implemented for the study. Participants spin the menu by rotating their thumb on the controller touchpad.

4 FORMATIVE: PILOT AND "IN THE WILD" STUDIES

Deploying a novel technical intervention in a new environment typically requires careful customization [38]. Prior to actually evaluating the utility of IA for economic analysis, we thus conducted a month-long formative study that included a pilot study (2 weeks) and an "in-the-wild" deployment (3 weeks). We opted to use the ImAxes platform [11] for immersive multidimensional visualization as our starting point; see the next section for details.

An added benefit of this formative approach is that it allowed us to continuously iterate on the design based on results from the user sessions as they occurred throughout the duration of the study. Participants were updated on notable changes to the system as they occurred and were asked to engage in additional tutorial, challenge, exploration, and interview activities following each major change to the system.

4.1 The ImAxes System

ImAxes [12] is an IA system based on the concept of *embodied axes* to let users build data views in a 3D virtual environment. Each axis corresponds to a dimension in a multivariate dataset. Users define visualizations by positioning axes in the 3D space, a spatial grammar producing specific visualizations based on their layout.

The basic operations consist of combining two or three orthogonal axes, which produces 2D or 3D scatterplots, respectively. Axes arranged in parallel to each other yield a parallel coordinates plot [19]. More advanced operations consist of stacking axes at the extremities of the axes of an existing scatterplot, which extends 2D and 3D scatterplots to scatterplot matrices. ImAxes also uses the proximity between visualizations to create linked 2D and 3D scatterplots.

4.2 Pilot Study

During the initial pilot study, we invited 6 participants to use the ImAxes platform for hour-long individual sessions with ImAxes left unmodified from its previous incarnation apart from the inclusion of an embedded tutorial. The dataset used during the pilot was the classic cars dataset [14]. The purpose of the pilot study was to: 1) identify the new features to add to ImAxes, 2) calibrate our data collection mechanisms, and 3) determine the datasets participants wanted to view.

4.3 "In-the-Wild" Study

After having established a working baseline system, we launched an "in-the-wild" formative study where the equipment stood available for a full three weeks for anyone to use at their own discretion. The author embedded at the agency advertised the study via the agency intranet, encouraging interested volunteers to bring their own datasets to explore. The room was kept unlocked, basic documentation was made available in the room, and the software configured to allow new participants to sign in and load their own data. However, while no experimenter was present during these sessions, IRB regulations required us to collect signatures of informed consent from volunteer participants. This allowed us to record video and telemetry whenever the equipment was in use. Similar to the pilot, the purpose of this study was to collect data for how to customize the system for an economist audience.

A total of six participants were engaged in this formative study (all provided signatures of informed consent; no unauthorized person used the tool). They logged a total of 3.8 hours of use in ImAxes during this phase. Figure 2 outlines the significant findings from our review of the logged data: this includes several observations that lead to refinements, as well as direct feature requests by the participants.

Throughout the three weeks of the formative study, we rolled out new features as soon as they were implemented, essentially using the field deployment as a "living laboratory."

4.4 Improvements to ImAxes

The original ImAxes system lacked many features necessary for an economics setting, including some that aid users regardless of domain background. We thus extended the system with additional features to support general use improvements to visual exploration and analysis of data based on feedback from economists. Below we list the main features added (labels refer to Figure 2).
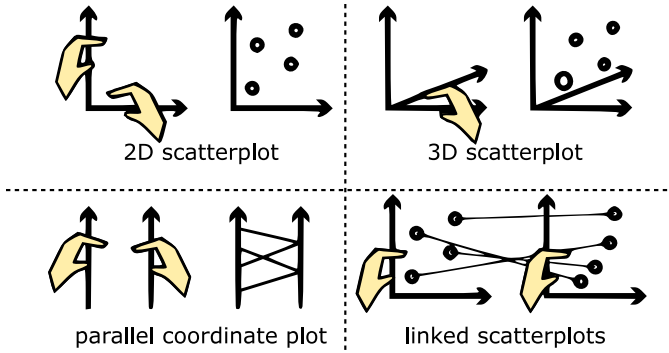
Fig. 5: Examples of using ImAxes to create different visualizations.

▷ **DR1: Tooltip (Details-on-demand).** We implemented details-on-demand as a tooltip for 2D and 3D visualizations using a pointer metaphor (Figure 3). By pressing a button on the controller and pointing in the direction of a 2D visualization, the data values of the nearest point are shown in a 2.5D box with a leader attached to the point. To obtain details-on-demand in a 3D scatterplot, a small pointer sphere is attached to the VR controller that can probe the nearest values.

▷ **DR2: Time-series data.** The original ImAxes supported only scatterplots and parallel coordinates plots. Since many of our users wanted to explore time-series data, we added line graphs as well.

▷ **DR3: Visualization design menu.** We added a simple menu control panel attached to the VR controller. This allows users to remap data dimensions to axes, create and bind a gradient colour to a continuous variable, and map the size of the points or lines to a data attribute.

▷ **DR5: Add mountain range backdrop.** Participants disliked the space's flat, sharp horizon and featureless terrain, causing us to add a mountainous landscape in the distance.

▷ **DR6: Axis selection.** Vanilla ImAxes used a *shelf metaphor* for selecting axes, where data axes were arranged in rows like books on a bookshelf.[1] While this metaphor is easy to understand, it does not scale with the number of axes and requires a large amount of locomotion (walking or teleporting). Our first solution, a "Rolodex" (DR4), was poorly received. Instead we implemented a rotational menu based on a "Lazy Susan" metaphor. The menu can be rotated like a Lazy Susan via the controller touchpad. An axis is selected by pulling it out of the Lazy Susan menu. Thus, in the end, there was no shelf.

▷ **DR7: Grouped selection.** Based on user feedback, we implemented a group selection mechanism that allows the user to move a linked group of visualizations instead of a single one. This enables the user to arrange visualizations around them without breaking links.

5 SUMMATIVE: CASE STUDIES IN ECONOMICS

To understand the utility of IA for professional analysts and data scientists, we conducted a contextual inquiry using our ImAxes tool in case studies involving participants from one of several bureaus of the U.S. federal government. This part of the study was split into two phases: Summative 1 (S1) and Summative 2 (S2). participants during S1 used a version of ImAxes that was slightly different in S2 (Figure 2). We report on both below, highlighting differences when needed. Unfortunately, a software error precluded collection of axis position data from S1. Other data was collected from both groups.

5.1 Participants

We recruited twelve participants (six in each phase) with expertise in economics, statistics, and data science. The participants were all employees at a U.S. federal agency with job descriptions that include data analysis, all with 2–12 years of experience ($M = 6.21, SD = 3.93$) and graduate degrees in economics or related fields (Table 1). They had significant experience in using data analysis tools in their daily work (Table 2). While outside the scope of this study, the typical workflow

[1] We need axes, lots of axes. (https://youtu.be/5oZi-wYarDs)

Table 1: Phases and datasets for summative participants.

#	Job Title	Yrs exp	Education	Phase*	Dataset†
1	Economist	12	M.A., Economics	S1	D1+D2
2	Economist	2	M.A., Economics	S1	D3+D4
3	Economist	9	Ph.D., Economics	S1	D1+D5
4	Economist	6	M.A., Economics	S1	D2
5	Economist	4	M.A., Economics	S1	D6
6	Econ spec.	8	M.A., Int. Business	S1	D6
7	Economist	2	M.A., Econ/Public Policy	F, S2	D7
8	Economist	5	M.A., Public Policy	S2	D8
9	Statistician	3	M.S., Statistics/Math	P, S2	D7
10	Economist	9	Ph.D., Economics	S2	D8
11	Economist	13	M.A., Economics	S2	D8
12	Economist	3	Ph.D., Economics	S2	D8

* P = pilot, F = formative ("in the wild"), S1/2 = summative 1/2
†Dataset labels in Table 3.

in government and industry data analysis is described by Batch and Elmqvist [5] and Kandel et al. [20]. Six participants had used VR previously, and five participants routinely played video games (1+ hr/wk).

Table 2: Count and context of participant use of specific tools.

Environment/Language	Ever	Work
Graphic analytical env. (e.g., Tableau, Excel)	12	11
Statistical lang. (e.g., Stata, R, SAS, Julia)	11	10
DBMS (e.g., SQL, PostGres, dBase)	8	6
Econometric DBMS (e.g., FAME, Aremos)	5	5
Markdown/doc-creation (e.g., HTML, LaTeX)	6	4
Object-oriented lang. (e.g., Python, JS)	7	3
Imperative lang. (e.g., FORTRAN, Pascal)	1	1

5.2 Procedure

Our study consisted of several stages: preparation, tutorial, exploration, presentation, and post-session interview.

▷ **Preparation**: Before participants even appeared at the study session, we asked them to send suitable datasets (Table 3) that we could integrate into ImAxes prior to the study. Some of these datasets caused us to make changes to the tool itself, such as the axis selection metaphor, as described in Section 4.4:DR6, which we implemented to accommodate a larger number of variables than practical in the shelf layout.

Table 3: Datasets used by summative participants.

#	Dataset Name
D1	Compensation by State and Industry
D2	Nominal PCE by State and Industry
D3	International Trade: Services
D4	U.S. Military Spending
D5	BLS Consumer Price Index
D6	National PCE Price Indexes
D7	Blended Health Care Satellite Accnt/Capita Exp. Index
D8	Nominal PCE by State

▷ **Instructional Tutorial** (10 mins): We began by familiarizing the users with ImAxes via a tutorial embedded in the system. Pre-recorded interactions were played, and the user was prompted to follow along to learn how to use the tool. During this stage, the affordances of a single axis were exhaustively demonstrated before moving on to two axes, then three, and finally SPLOMs and parallel coordinate plots. After each feature was demonstrated, we asked participants to use that

feature in a sample dataset. Before finishing the training, participants were encouraged to freely explore the sample dataset while verbalizing their thought process using a think-aloud protocol.

▸ **Exploration** (30 mins): Now participants were set free to explore their own dataset on their own. Exploration was structured as a sequence of iterations, each no less than five minutes, and started with giving the participant the option of introducing a new dataset if desired. For each iteration, the researcher prompted the participant to maintain the think-aloud protocol, and would gently inquire about their motivations throughout the duration. The goal of each iteration was to generate at least one insight and corresponding visualization. Participants were told that they would be expected to present their findings, and were regularly updated on remaining time.

▸ **Presentation** (30 mins): Finally, the participant was asked narrate their findings as if they were presenting their analysis to an external party (the experimenter). The participant was reminded that the experimenter could see what they saw on a monitor, and was asked to create at least one distinct visualization for each point in their narrative. They could use speech, gestures, and ImAxes itself to tell their story.

▸ **Post-Session Interview**: Immediately after the exploration activity, the researcher and participant engaged in a brief, semi-structured interview and survey to (a) validate the researcher's understanding of the user's motivations for their actions during the exploration activity, and (b) Evaluate the user's sentiment regarding the existing iteration of the implementation, including features that they felt were lacking.

5.3 Predictions

In this preregistered study,[2] we made several predictions on results prior to conducting the study. We organize these predictions into the stage they refer to: exploration (*E*), presentation (*P*), and all (*A*).

E1 Participants will arrange the views egocentrically around themselves. *Motivation:* For individual work, it is more efficient to use local space around yourself.

E1.1 Participants will tend to arrange their views at chest level. *Motivation:* Participants have no specific VR training, and will thus likely not utilize the 3D environment to the fullest.

E1.2 Participants will arrange their views within easy reach of the center of the space. *Motivation:* The small space that the study is conducted in will not permit significant physical navigation.

E2 Participants will build many ephemeral visualizations that they quickly discard. *Motivation:* ImAxes supports exploration by creating transient and new visualizations through brushing.

P1 Participants will arrange the views in an exocentric way. *Motivation:* During presentation, it makes sense to more carefully arrange the views, e.g., in a gallery or sequence.

P1.1 Participants will arrange the views in a chronological order w.r.t. to their presentation. *Motivation:* The intelligent use of physical space can help streamline a narrative.

P2 Participants will build more complex visualizations in the presentation stage than the explore stage. *Motivation:* Presentation involves creating a linear, coherent, and comprehensive narrative. Care can thus be spent on crafting complex visualizations. *Note:* This prediction was not part of our preregistration.

A1 Participants will prefer basic visual representations (scatterplots, linegraphs, maps), and avoid more complex ones (parallel coordinates, scatterplot matrices). *Motivation:* These more complex representations are not commonplace in real-world data analysis.

A1.1 Participants will avoid using 3D representations (such as 3D scatterplots or surfaces). *Motivation:* Our participants have no VR training and are accustomed to 2D displays in their work.

A2 Participants will utilize the physical space to structure their work. *Motivation:* Physical space can be used to support specific tasks, e.g., to simplify choice, perception, and computation [22].

A2.1 Participants will group views in space based on their logical relationships. *Motivation:* Views that belong together should be grouped in physical proximity.

A2.2 Participants prefer interacting with objects at a near distance than those at a far distance. *Motivation:* Near objects require no physical navigation to access, and ImAxes does not support a reliable distance interaction technique.

A3 Participants will report typical perceptual and cognitive effects of VR on their performance and perception. *Motivation:* Even if ImAxes depicts an abstract data analysis setting, it is subject to the same strengths and weaknesses as other VR applications.

A3.1 Participants will report a high level of engagement. *Motivation:* VR is commonly associated with high engagement because of realism and low indirection.

A3.2 Participants will report a high level of presence. *Motivation:* VR is commonly associated with high presence because of the low indirection, natural interaction, proprioception, and the perception of physical space.

A3.3 Participants will report fatigue from physical navigation and interaction. *Motivation:* The use of gross body motor controls to navigate the virtual environment and interact with its objects will require significant effort by the participants.

A3.4 Participants will suffer from reduced legibility of text in the 3D environment. *Motivation:* HMDs have a significantly lower resolution than typical monitors, and labels in ImAxes are 3D and thus subject to distance and orientation concerns.

A3.5 Participants will suffer from the challenge of using VR wands to interact with virtual 3D objects. *Motivation:* While more direct than using a mouse and keyboard, the HTC Vive controllers still do not allow for hand and finger interaction.

A4 Participants will encounter significant navigation and interaction hurdles due to a lack of VR expertise. *Motivation:* Our participant pool has no specific VR training, and will thus be challenged by 3D navigation and interaction concerns.

A4.1 Participants with 3D computer gaming experience will be less hindered by lack of VR training. *Motivation:* 3D gaming experience will help people interact more efficiently.

6 CASE STUDY RESULTS

Table 1 reviews the participants and their datasets. Below we discuss a representative use case derived from the experiment. We then present the performance and subjective results.

6.1 Representative Use Case

The following scenario is a pastiche based on our observations of participants as they explored and presented insights from their macroeconomic data. It is not a description of a single user session; rather, it is a collection of real observations from multiple sessions organized into a representative, narrative summary. In other words, unlike the scenario in the introduction, it is not fictional; these events all happened. The scenario begins with our economist "Sasha" loading their regional personal consumer expenditures dataset into ImAxes. Sasha has just found this dataset from a public source and wants to explore the 2007–2009 Great Recession's effect on trends in consumer expenditures.

❧   ❧

SASHA dons their VR headset, launches ImAxes, and grabs three axes from the Lazy Susan. They build a 3D scatterplot of TimePeriod × Goods × GeoFips (states) by first holding TimePeriod and Goods orthogonal to each other, then placing GeoFips orthogonal to the scatterplot's origin. They orient the visualization so that they are looking down the temporal axis, leveraging the depth perception afforded by VR to provide a view of the states where the goods have trended higher over time. Using this view, they activate the details on demand using the controllers for these states to obtain numeric values of points along the axes.

Sasha then flips the view so that they are looking at TimePeriod from the side, and points out the general upward trend in total consumer spending for all commodities over all time periods except the Great Recession around 2009. They create a 2D scatterplot of gasoline expenditures over time, noting that the trend is less stationary (i.e., has greater variance over time) in that particular commodity than in others.

Sasha creates a 3D scatterplot of Food Services × TimePeriod × Off Premises Food and Drink. Grabbing another Time Period axis, they switch from a 3D scatterplot to two separate 2D scatterplots, which they stack on top of each other. They observe that there is a switch from spending on restaurants (Food Services) toward spending more on groceries (Off Premises Food and Drink) during the Great Recession.

Once they have constructed all of charts they intend to discuss with their colleagues, they arrange them in the space in a linear order from left to right roughly corresponding to the narrative order they plan on following, a little like a museum or gallery of artifacts. As they discuss each point, starting with the most aggregate commodity bundles and drilling down into more detailed commodities, they dynamically interact with the visualization with one or both hands, shifting it for a different viewing angle with one hand and calling the tooltip with the other hand to give their expert audience the detail they would otherwise demand. When they are done discussing the points related to one visualization, they walk to the right to begin their next talking point, until they have run through all of the economic trends they wish to discuss.

☙ ❧

6.2 Explore Stage

Participants spent between 4 and 10 minutes (M =4:33, SD =1:50) in the explore stage. All participants would begin the stage by facing the Lazy Susan within arm's reach, and would rotate it until they found an axis they recognized from which they could start exploring the data. Participants would then often rotate their body away from the Lazy Susan to create a work space by building basic 2D and 3D scatterplots. Figure 6 shows that most participants stayed in one place and arranged views egocentrically (**E1**). However, none utilized the full 360° space.

This behavior of recycling the views and axes in their workspace instead of physically moving to a new workspace also supports prediction **E1.2** (participants would arrange their views within easy reach).
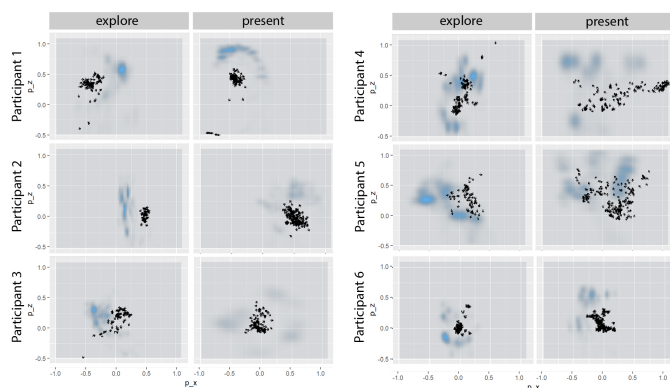
Fig. 6: Heatmaps of axis interaction in S2 (top-down). Participant position and view direction is represented by a direction arrow.

To examine prediction **E1.1**—that participants would arrange views at roughly chest level—we studied interaction patterns w.r.t. height. Since our tracking data only includes headset position, we estimate chest height to be approximately 30 cm below this position. Figure 7
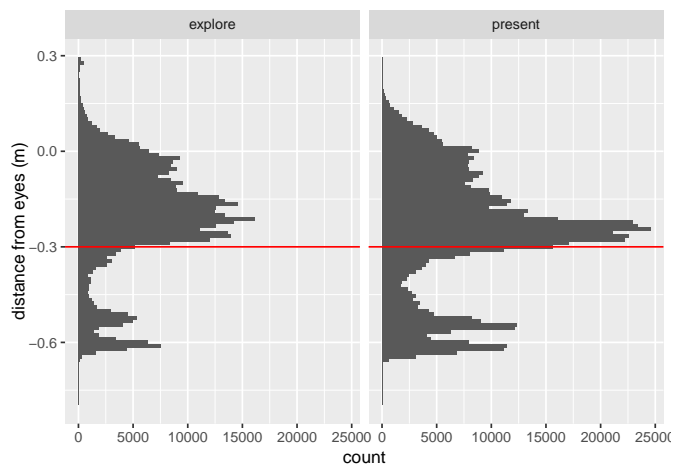
Fig. 7: Histogram showing the vertical distance of participant interactions with axes relative to their eye level. Eye level is at 0, and the approximate chest level is represented by the red line.

shows a histogram of these relative interactions. Interaction above eye level often occurred when building scatterplot matrices, highlighting a physical limitation of the ImAxes systems (that a user must be able to reach the ends of a scatterplot matrix). While this limitation is somewhat mitigated by design refinement **DR7** (grouping mechanism), these observations suggest that the issue is still present.

Participants would often discard axes and visualizations while exploring the data, maintaining only one to two visualizations at a time (supporting **E2**). Essentially, participants were recycling their views and continuously cleaning their space. Furthermore, we observed that certain types of visualization would be more transient than others. Notably, linked visualizations, whether between two axis or between an axis and a scatterplot, were created and used more than any other type of visualization, but the majority existed for less than five seconds.

6.3 Presentation Stage

Participants spent between 7 and 11 minutes each (M =6:30, SD =2:30) during the presentation stage. Most participants chose to organize their views in either a linear or semi-circular layout. For example, Participants 4 and 5 placed a series of visualizations in a left to right "narrative order" (Figure 6). This somewhat supports prediction **P1** (participants will arrange the views in an exocentric way). However, as can be seen in the "present" columns of Figure 6, these arrangements were not strictly exocentric, but remained egocentric (undermining **P1**). We belatedly realized that since the experimenter—the intended audience of the presentation—viewed the 3D space through the eyes of the participant, there was no incentive for the participant to organize the space in an exocentric fashion. However, we did find support for views being arranged in chronological order (**P1.1**); Figure 1 shows snapshots of several final view layouts.

We predicted (**P2**) that participants would build more complex visualizations during the presentation stage as they would spend time to carefully craft a meaningful visualization. This is also supported by our data; Table 4 indicates that most scatterplot matrices were used in the presentation stage. All participants except Participant 1 created a scatterplot matrices while preparing for presentation stage; however, only Participant 3 actually used a scatterplot matrix when presenting their data. Five of the six participants explored the data using parallel coordinate plots. However, it is worth noting that during the presentation stage, only Participant 3 used a parallel coordinate plot.

6.4 All Stages

We captured view creation events for 2D and 3D scatterplots, SPLOMs, and linked views. These events are summarized in Table 4. Contrary to prediction **A1** (that participants would avoid complex vi-

Table 4: Count of view creations per participant, split into exploration (**E**) and presentation (**P**) stages.

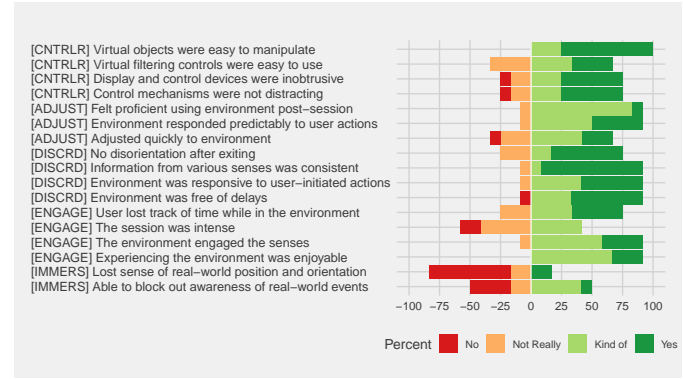| Participant	2D Scatterplot		3D Scatterplot		SPLOM		Link	
	E	**P**	**E**	**P**	**E**	**P**	**E**	**P**
1	8	17	10	2	29	-	111	113
2	9	33	-	12	5	5	116	201
3	27	74	6	16	-	26	103	4136
4	7	43	1	11	8	18	123	516
5	5	58	2	15	6	22	31	195
6	3	49	-	23	2	35	527	2866

Fig. 8: Subjective ratings from exit survey. Subfactors include: controller ease of use [CNTRLR], adjustment to environment [ADJUST], perceived immersion [IMMERS], user engagement [ENGAGE], and avoidance of sensory discord [DISCRD].

sualizations), all participants (except P3) experimented with creating scatterplot matrices during exploration. The majority of these scatterplot matrices involved adding a third axis to an existing 2D scatterplot in order to see the relationship between two variables and on a common axis (such as a time-series axis).

All participants except P2 and P6 created 3D scatterplots during the explore stage. However, all participants used 3D scatterplots during their presentation stage. Notably, P2 and P6 used 3D scatterplots exclusively during the presentation stage, and P5 used three 3D scatterplots and a 2D scatterplot during the presentation stage. This result ran counter to prediction **A1.1** (participants would avoid creating 3D visualizations). One participant commented that they felt they may as well use 3D scatterplots and other kinds of visualizations as they were exploring data in VR, saying "I wanted to create more graphs of different types, [especially for] my presentation."

We found only weak support for **A2**; during the explore stage, participants would merely choose the nearest open space for creating new views, i.e., not using an organizing principle. Only in the explore stage were they more conscious of structuring the space; more specifically, as noted in our observations supporting **P1.1**, chronology was a common such organizing principle (also partially supporting **A2.1**). We also noted that many undertook a "curation" stage where they would select views that should be included in the presentation, and move them to a designated area.

When considering **A2.2**, we expected participants to minimize their walking, relying instead on rotating their viewpoint. We found that participants walked less during the explore stage compared to the presentation stage. We ran a paired sample t-test to compare the movement per minute of the explore and present stages. The present stage ($M = 8.10$m, $SD = 2.18$m) had significantly more movement than the explore stage ($M = 5.41$m, $SD = 1.8$m); $t(5) = -3.456, p = 0.018$ (see chart). One participant commented that "When I start thinking of myself as a visual focal point rather than thinking of myself as being surrounded by [vertical] boards, viewing the environment became easier and I felt comfortable using more of the space." We did not find a significant difference in head rotations per minute between the present stage ($M = 6671.86, SD = 5397.94$) and the explore stage ($M = 2571.38, SD = 1321.33$); $t(5) = -2.277, p = 0.072$.

6.5 Self-reported Perceptual and Cognitive Effects

Even if ImAxes depicts an abstract data analysis setting, it is subject to the same strengths and weaknesses as a general virtual environment; Figure 8 shows self-reported perceptual and cognitive effects similar to typical such environments (**A3**). According to the figure, participants reported high scores for perceived engagement (**A3.1**), rating the experience as enjoyable and engaging the senses.

We expected participants to report a high level of presence (**A3.2**) using the system. Supporting this prediction, the survey responses (Figure 8) show that participants felt that they were interacting in a natural environment (100% described the environment as being realistic and generally feeling natural, 83% felt moving around was natural). One user described the experience as being somewhat like "*being in the Mojave Desert.*" Several reported that they lost track of time while in the environment. However, full presence may not always be ideal; two thirds of participants reported that their exploration of the data was not intense, and several users pointed out that the sound of the researcher's voice improved their sense of orientation in the real room.

As for increased fatigue level (**A3.3**), we were not able to find support for this prediction. In fact, as our discussion for **A2.2** shows, physical navigation actually *increased* for the presentation stage (which followed exploration), suggesting that fatigue was not a factor. Furthermore, the nausea level was low, which is another indication that participants were not fatigued by the end of the study. We made several predictions related to the challenges that an immersive VR system could potentially introduce. We predicted that participants would suffer from reduced text legibility (**A3.4**). However, the reported Likert scores for the ability to examine closely and obtain details from objects in the environment was high (Figure 8), which undermines this prediction. We also expected that participants would suffer from the impreciseness of the VR wand controllers (**A3.5**). Again to the contrary, the Likert scores indicate that actually participants were able to effectively interact with the 3D virtual objects in ImAxes. By and large, one of the things the participants stated liking most about the environment was that visualizations were very fast and easy (or intuitive) to create relative to their traditional 2D working environment. However, there was a more even split in the participants in regards to wearing the physical VR headset, with one participant reporting "*I don't really like wearing a headset. It's cool to look at things in 3D, but it doesn't really add enough value. However, I also don't usually create visualizations in general during my analyses.*"

Finally, we made two predictions in regards to VR experience; that participants with a lack of VR experience would encounter significant navigation issues (**A4**), and that gaming experience would mitigate a lack of VR experience (**A4.1**). Based on differences in survey responses for users with VR experience versus those without, we find support for the first of these predictions, but not for the second. In fact, participants without VR experience who regularly play computer games for more than an hour in a week reported having more difficulty with the controls and had a more difficult time examining objects in the environment than those who were not regular gamers.

6.6 Qualitative Feedback

Beyond interaction and visualization requests, participants provided several insightful comments. Whiteboard analogies were commonplace: "*It feels like I'm surrounded by whiteboards,*" said one user; another, after arranging axes around himself in a semi-cylinder shape,

described it as feeling like a "*wraparound whiteboard.*"

When asked how ImAxes compared to their traditional desktop display, participants had a range of responses; 58.3% reported feeling more engaged in the problem while in ImAxes than while in their traditional environment (**A3.2**). The most common draw participants felt to the environment was that creating visualizations was easier and faster in ImAxes than in their typical environments. In general, participants said they might be able to use ImAxes for preparing presentations, reports, and video communications, or for exploratory analysis data validation. One participant responded that they could use it to detect errors during the monthly multi-stage process of reviewing economic indicator estimates prior to publication. Another said that their indicator estimation process involves multilateral aggregation for price indices, and ImAxes could be useful for exploratory analysis during that process. One user noted that export for 2D display presentation would be particularly helpful for the purpose of creating reports.

Several participants said that a major barrier to wanting to use ImAxes on the job is that VR is inconvenient for the purpose of the type of work they perform, which typically involves programming and switching between multiple environments. Said one participant, "*VR seems more oriented toward real-time demonstrations, which is great, but that's not useful for [the participant's] analytical process, which involves long periods of exploration and evaluation switching between tabular views, charts, modeling, and programming.*"

While we were able to implement some changes between our formative phase and our summative phase, there were some changes that were not practical to implement during the span of this study; some of these might be considered applicable for general use, while others are more economics domain-specific. One participant, who was not interested in using ImAxes on the job, said "*it would be sick [sic] if I could click something and see the full hierarchy of categories in the data.*" The absence of this feature wound up being the primary reason for their recalcitrance. Other features this particular participant wanted to see included the ability to run regressions, a group-by mechanism, extra-grammatical filtering mechanisms for building views, and simple computational tasks. Like this participant, several other participants during the summative and earlier phases of the study suggested the inclusion of matrix and column-wise algebraic operations. A number of participants also requested linear modeling operations and views of multicollinearity, which they noted as being particularly relevant for hedonic modeling. Another common request was that we extend ImAxes as a tool specializing in outlier detection. Economists typically evaluate time series; while our addition of a line mark connecting scatterplot points was one change we did implement to accommodate this activity, participants regularly reused time period axes, and the option of having a convenient "favorite axes" quick-access area was requested by multiple users. Finally, one participant strongly suggested the addition of a Markov Chain Monte Carlo simulation, stating that it is "what everyone is doing now" in econometric modeling.

7 CONCLUSION AND DISCUSSION

We have presented a design study on the use of immersive analytics for economics analysis. The entirety of the work was conducted in the field. Our multi-phase design process involved an initial requirements analysis and pilot study on the use of the ImAxes immersive analytics toolkit for economics, a field deployment during which the tool design was iteratively refined, and several in-depth case studies with professional analysts. Our findings illuminate several unexpected insights about the nature of immersive analytics for experts.

We use VR headsets for our evaluation as opposed to Augmented Reality (AR) devices such as Magic Leap or HoloLens since current commercially available VR devices offer vastly superior field of view and resolution compared to their AR counterparts. AR headsets have advantages for collaborative analysis by permitting a clear view of colleagues, or for situated analytics [48] where data can be overlaid on objects in the world, but these are not the focus of this paper; hence VR is the best fit for this purpose. As AR technology improves there is a real possibility that at some point immersive headset technologies will make screens obsolete, such as when headsets can render a 2D

display at foveal resolution completely covering the wearer's field of view anywhere in the space around them. Regardless, AR and VR alike both require knowing how to make the most of immersive spaces for data visualization in real-world applications such as economics.

We were surprised that many of our predictions found no support in the collected data. For example, we noticed few effects of fatigue (**A3.3**), legibility was not a clear concern (**A3.4**), and even participants with little gaming and/or VR experience were able to use our tool efficiently (**A3.5**). Some of these findings can be easily explained—e.g., that the lack of an exocentric layout likely happens because presenters actually view the environment through the eyes of the participant (**P1**)—but others are more unexpected. Most of the time, while aptly highlighting our lack of knowledge, these contrary results are actually in favor of IA; for example, participants did actually use advanced visualizations (**A1**), not merely sticking to scatterplots.

On the surface, this finding disappointingly does not extend to the intelligent use of space (**A2**), as participants in the explore stage merely picked the closest free space to put new visualizations. However, when viewed through the sensemaking loop [33], this makes more sense as one of its early phases involves placing potentially relevant information in a so-called "shoebox." When foraging for information, analysts typically do not have time to worry about structure, similar to how the purpose of sketching for artists is to generate new ideas rather than fixate on existing ones. Only in a secondary curation step in our study would participants evaluate these views and organize them into a designated area in the environment (the "evidence file" in the sensemaking loop). Incidentally, Andrews [1] noted similar observations, referring to them as "evidence marshalling," often having the same chronological organizing principle (**A2.1**) as our findings.

Still, it is clear that participants did not use the full available 3D space of the analysis environment to its full potential. Telemetry data and video recordings showed participants mostly stayed in place and merely used the space directly in front of them. We did see participants making better use of the space in the presentation stage. Reasons for this may include the cramped confines of our experimental space, the interactions needed to move visualizations, and no automatic layout control. This points to the need for system support, such as constraints and organization frameworks, to help users organize their spaces as has previously been done for 2D GUI tools [27, 31].

That many of our—in retrospect pessimistic—predictions about the drawbacks of the virtual environment were not supported is worth unpacking. One reason may be the high presence and engagement levels reported, leaving participants willing to simply overlook minor usability concerns. The novelty factor may also be working in our favor and produce goodwill towards the tool.[3] Finally, perhaps the natural interaction metaphors in ImAxes simply aided participants in quickly learning the system and exploring their data. In fact, we were surprised by the level of interest in using ImAxes in the workplace. There were a number of domain-specific tasks and requests which we were unable to fully accommodate in the scope of this study—MCMC simulation, linear modeling and views of multicollinearity for hedonics, a suite of features for outlier detection—as well as the more generally-applicable requests for matrix algebraic operations, quick-access-axes for "favorite" axes (like time series period indices), and integration with hierarchical views of the data. While these changes were not practical to implement in this study, we view them as low-hanging fruit for future work extending immersive implementations either for general analytical tasks or more specialized economic analysis ones.

---

[3]Do not try and bend the spoon. That's impossible. Instead, only try to realize the truth... there is no spoon. Then you'll see that it is not the spoon that bends, it is only yourself. (https://youtu.be/uAXtO5dMqEI)

REFERENCES

[1] C. Andrews, A. Endert, and C. North. Space to think: large high-resolution displays for sensemaking. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 55–64. ACM, 2010. doi: 10.1145/1753326.1753336

[2] R. Arias-Hernandez, L. T. Kaastra, T. M. Green, and B. Fisher. Pair analytics: Capturing reasoning processes in collaborative visual analytics. In *Proceedings of the Hawaii International Conference on System Sciences*, pp. 1–10. IEEE, 2011. doi: 10.1109/HICSS.2011.339

[3] R. Ball, C. North, and D. A. Bowman. Move to improve: promoting physical navigation to increase user performance with large displays. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 191–200. ACM, New York, NY, USA, 2007. doi: 10.1145/1240624.1240656

[4] D. Banakou, R. Groten, and M. Slater. Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences*, 110(31):12846–12851, 2013. doi: 10.1073/pnas.1306779110

[5] A. Batch and N. Elmqvist. The interactive visualization gap in initial exploratory data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):278–287, Jan. 2018. doi: 10.1109/TVCG.2017.2743990

[6] H. R. Beyer and K. Holtzblatt. *Contextual Design: Defining Customer-Centered Systems*. Morgan Kaufmann, San Francisco, CA, USA, 1997.

[7] D. A. Bowman, C. North, J. Chen, N. F. Polys, P. S. Pyla, and U. Yilmaz. Information-rich virtual environments: theory, tools, and research agenda. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, pp. 81–90. ACM, New York, NY, USA, 2003. doi: 10.1145/1008653.1008669

[8] S. Butscher, S. Hubenschmid, J. Müller, J. Fuchs, and H. Reiterer. Clusters, trends, and outliers: How immersive technologies can facilitate the collaborative analysis of multidimensional data. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 90:1–90:12. ACM, New York, NY, USA, 2018. doi: 10.1145/3173574.3173664

[9] T. Chandler, M. Cordeil, T. Czauderna, T. Dwyer, J. Glowacki, C. Goncu, M. Klapperstueck, K. Klein, F. Schreiber, and E. Wilson. Immersive analytics. In *Proceedings of the International Symposium on Big Data Visual Analytics*, pp. 1–8. IEEE, Piscataway, NJ, USA, 2015. doi: 10.1109/BDVA.2015.7314296

[10] A. Clark and D. Chalmers. The extended mind. *Analysis*, 58(1):7–19, 1998. doi: 10.1093/analys/58.1.7

[11] 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 ACM Symposium on User Interface Software and Technology*, pp. 71–83. ACM, New York, NY, USA, 2017. doi: 10.1145/3126594.3126613

[12] M. Cordeil, T. Dwyer, K. Klein, B. Laha, K. Marriott, and B. H. Thomas. Immersive collaborative analysis of network connectivity: CAVE-style or head-mounted display? *IEEE Transactions on Visualization and Computer Graphics*, 23(1):441–450, 2017. doi: 10.1109/TVCG.2016.2599107

[13] A. Druin. Cooperative inquiry: Developing new technologies for children with children. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 592–599. ACM, New York, NY, USA, 1999. doi: 10.1145/302979.303166

[14] D. Dua and C. Graff. UCI machine learning repository. http://archive.ics.uci.edu/ml, 2017.

[15] T. Dwyer, K. Marriott, T. Isenberg, K. Klein, N. Riche, F. Schreiber, W. Stuerzlinger, and B. H. Thomas. Immersive analytics: An introduction. In *Immersive Analytics*, vol. 11190 of *Lecture Notes in Computer Science*, pp. 1–23. Springer, 2018. doi: 10.1007/978-3-030-01388-2_1

[16] G. Fontaine. The experience of a sense of presence in intercultural and international encounters. *Presence: Teleoperators and Virtual Environments*, 1(4):482–490, Jan. 1992. doi: 10.1162/pres.1992.1.4.482

[17] C. Hurter, N. Riche, S. 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

[18] E. Hutchins. *Cognition in the Wild*. MIT Press, Cambridge, MA, USA, 1995.

[19] A. Inselberg and B. Dimsdale. Parallel coordinates: a tool for visualizing multi-dimensional geometry. In *Proceedings of the IEEE Conference on Visualization*, pp. 361–378. IEEE, Piscataway, NJ, USA, 1990. doi: 10.1109/VISUAL.1990.146402

[20] S. Kandel, A. Paepcke, J. M. Hellerstein, and J. Heer. Enterprise data analysis and visualization: An interview study. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2917–2926, Dec 2012. doi: 10.1109/TVCG.2012.219

[21] K. Kilteni, R. Groten, and M. Slater. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387, 2012. doi: 10.1162/PRES_a_00124

[22] D. Kirsh. The intelligent use of space. *Artificial Intelligence*, 73(1–2):31–68, 1995. doi: 10.1016/0004-3702(94)00017-U

[23] D. Kirsh and P. Maglio. On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18(4):513–549, 1994. doi: 10.1207/s15516709cog1804_1

[24] D. Koller, P. Lindstrom, W. Ribarsky, L. F. Hodges, N. Faust, and G. Turner. Virtual GIS: a real-time 3D geographic information system. In *Proceedings of the IEEE Conference on Visualization*, pp. 94–100. IEEE, Piscataway, NJ, USA, 1995. doi: 10.1109/VISUAL.1995.480800

[25] B. Laha, D. A. Bowman, and J. J. Socha. Effects of VR system fidelity on analyzing isosurface visualization of volume datasets. *IEEE Transactions on Visualization & Computer Graphics*, 20(4):513–522, Apr. 2014. doi: 10.1109/TVCG.2014.20

[26] J. Lessiter, J. Freeman, E. Keogh, and J. Davidoff. A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments*, 10(3):282–297, 2001. doi: 10.1162/105474601300343612

[27] M. A. Linton, J. M. Vlissides, and P. R. Calder. Composing user interfaces with InterViews. *Computer*, 22(2):8–22, 1989. doi: 10.1109/2.19829

[28] K. Marriott, F. Schreiber, T. Dwyer, K. Klein, N. H. Riche, T. Itoh, W. Stuerzlinger, and B. H. Thomas, eds. *Immersive Analytics*, vol. 11190 of *Lecture Notes in Computer Science*. Springer, 2018. doi: 10.1007/978-3-030-01388-2

[29] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Proceedings of Telemanipulator and Telepresence Technologies.*, vol. 2351, pp. 282–292. SPIE, 1995. doi: 10.1117/12.197321

[30] A. Mottelson and K. Hornbæk. Virtual reality studies outside the laboratory. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, pp. 9:1–9:10. ACM, New York, NY, USA, 2017. doi: 10.1145/3139131.3139141

[31] B. A. Myers, D. A. Giuse, R. B. Dannenberg, B. V. Zanden, D. S. Kosbie, E. Pervin, A. Mickish, and P. Marchal. Garnet: comprehensive support for graphical, highly interactive user interfaces. *Computer*, 23(11):71–85, Nov 1990. doi: 10.1109/2.60882

[32] D. A. Norman. Cognition in the head and in the world: An introduction to the special issue on situated action. *Cognitive Science*, 17(1):1–6, 1993. doi: 10.1207/s15516709cog1701_1

[33] P. Pirolli and S. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of the International Conference on Intelligence Analysis*, vol. 5, pp. 2–4, 2005.

[34] K. Reda, A. E. Johnson, M. E. Papka, and J. Leigh. Effects of display size and resolution on user behavior and insight acquisition in visual exploration. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 2759–2768. ACM, New York, NY, USA, 2015. doi: 10.1145/2702123.2702406

[35] W. Ribarsky, J. Bolter, A. O. den Bosch, and R. Van Teylingen. Visualization and analysis using virtual reality. *IEEE Computer Graphics and Applications*, 14(1):10–12, 1994. doi: 10.1109/38.250911

[36] G. Robertson, M. Czerwinski, K. Larson, Robbins, D. C., D. Thiel, and M. van Dantzich. Data mountain: Using spatial memory for document management. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, pp. 153–162. ACM, New York, NY, USA, 1998. doi: 10.1145/288392.288596

[37] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, 10(3):266–281, June 2001. doi: 10.1162/105474601300343603

[38] M. Sedlmair, M. D. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440, 2012. doi: 10.

1109/TVCG.2012.213

[39] L. Shapiro. *Embodied Cognition*. Routledge, New York, NY, USA, 2011.

[40] T. B. Sheridan. Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments*, 1(1):120–126, 1992. doi: 10.1162/pres.1992.1.1.120

[41] K. Shilton. This is an intervention: Foregrounding and operationalizing ethics during technology design. In K. D. Pimple, ed., *Emerging Pervasive Information and Communication Technologies: Ethical Challenges, Opportunities and Safeguards*, pp. 177–192. Springer, Dordrecht, May 2014. doi: 10.1007/978-94-007-6833-8_9

[42] F. M. Shipman III, C. C. Marshall, and T. P. Moran. Finding and using implicit structure in human-organized spatial layouts of information. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 346–353. ACM, New York, NY, USA, 1995. doi: 10.1145/223904.223949

[43] B. Shneiderman and C. Plaisant. Strategies for evaluating information visualization tools: multi-dimensional in-depth long-term case studies. In *Proceedings of the AVI Workshop on BEyond time and errors: novel evaluation methods for information visualization*, pp. 1–7. ACM, New York, NY, USA, 2006. doi: 10.1145/1168149.1168158

[44] M. Simpson, J. O. Wallgrün, A. Klippel, L. Yang, G. Garner, K. Keller, D. Oprean, and S. Bansal. Immersive analytics for multi-objective dynamic integrated climate-economy (DICE) models. In *Proceedings of the ACM Companion on Interactive Surfaces and Spaces*, pp. 99–105. ACM, New York, NY, USA, 2016. doi: 10.1145/3009939.3009955

[45] M. Slater, M. Usoh, and A. Steed. Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 3(2):130–144, Jan. 1994. doi: 10.1162/pres.1994.3.2.130

[46] A. Steed, S. Friston, M. M. López, J. Drummond, Y. Pan, and D. Swapp. An 'in the wild' experiment on presence and embodiment using consumer virtual reality equipment. *IEEE Transactions on Visualization and Computer Graphics*, 22(4):1406–1414, Apr. 2016. doi: 10.1109/TVCG.2016.2518135

[47] D. S. Tan, D. Gergle, P. G. Scupelli, and R. Pausch. With similar visual angles, larger displays improve performance on spatial tasks. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 217–224. ACM, New York, NY, USA, 2003. doi: 10.1145/642611.642650

[48] B. H. Thomas, G. F. Welch, P. Dragicevic, N. Elmqvist, P. Irani, Y. Jansen, D. Schmalstieg, A. Tabard, N. A. M. ElSayed, R. T. Smith, and W. Willett. Situated analytics. In K. Marriott, F. Schreiber, T. Dwyer, K. Klein, N. H. Riche, T. Itoh, W. Stuerzlinger, and B. H. Thomas, eds., *Immersive Analytics*, vol. 11190 of *Lecture Notes in Computer Science*, pp. 185–220. Springer, 2018. doi: 10.1007/978-3-030-01388-2_7

[49] A. van Dam, D. H. Laidlaw, and R. M. Simpson. Experiments in immersive Virtual Reality for scientific visualization. *Computers & Graphics*, 26(4):535–555, 2002. doi: 10.1016/S0097-8493(02)00113-9

[50] J. Vasconcelos-Raposo, M. Bessa, M. Melo, L. Barbosa, R. Rodrigues, C. M. Teixeira, L. Cabral, and A. A. Sousa. Adaptation and validation of the Igroup presence questionnaire (IPQ) in a Portuguese sample. *Presence: Teleoperators and Virtual Environments*, 25(3):191–203, 2016. doi: 10.1162/PRES_a_00261

[51] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240, 1998. doi: 10.1162/105474698565686

[52] D. Wixon, K. Holtzblatt, and S. Knox. Contextual design: An emergent view of system design. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 329–336. ACM, New York, NY, USA, 1990. doi: 10.1145/97243.97304