

# Comparative Glyph-Field Trajectory Analyses with an AR+Tablet Hybrid User Interface for Geospatial Analysis Tasks

S. Hu<sup>1</sup>  and D. Reilly<sup>1</sup> 

<sup>1</sup>Graphics and Experiential Media Lab, Faculty of Computer Science, Dalhousie University, Canada

## Abstract

Augmented reality (AR) supports large virtual display areas without the need for physical screens—affording more mobility to the user and displaying map-based data. Current head-worn AR devices have limited processing and rendering capabilities. Their hand-free input is imprecise. Hybrid interfaces, such as AR+tablet, can mitigate these limitations: the tablet can provide additional display fidelity in a region of interest and act as a precise input device. Used together, AR and a tablet support tasks that simultaneously require mobility and large area displays. However, more work is needed on such a system to understand the influence of glyph visualization techniques on glyph field scanning behaviours. Two glyph-based representations named Polyline and Mondrian were compared. Polyline is a shape-based technique known to be good for finding trends in desktop contexts. Mondrian is a colour-based technique. In theory, it is good for pre-attentive cursory exploration. Participants performed semi-naturalistic tasks based on geospatial linear regression. Polyline induced more scrolling on the tablet because participants wanted to examine glyphs more closely. Mondrian induced more gaze movement across the AR display region, but tasks could also affect gaze. We then discuss focus+context, and colourmap design.

## CCS Concepts

• **Human-centered computing** → **Mixed / augmented reality; Tablet computers; Visualization design and evaluation methods;**

## 1. Introduction

Advances in augmented reality (AR) and optical see-through head-worn displays (OST-HWD) allow us to have a room-sized display area with minimal cost to mobility [PNB\*21, WSS20]; permitting large 2D elements like maps and 3D content. However, AR displays typically do not have high display resolutions [ILSP21], and often do not facilitate precise input [FS91], despite advances in AR input methods (e.g., gaze and gesture-based [KC17]) [SBSPM21]. Combining AR with a tablet computer provides a surface with superior graphical fidelity and more accurate touch-based input [SGHV19]. Our work explores how a tablet+AR configuration supports scanning a room-sized glyph field. A glyph field is a large collection of glyphs arranged using positional data (e.g., arrangement based on latitude and longitude [FIBK17]), and a glyph is a visual data marker arranged to express data or information.

Understanding how the user scans a glyph field on an AR-extended tablet screen is critical for designing an optimal AR+tablet interface. We collected and analyzed the trajectory data (gaze, panning with tablet-based scroll, and position) generated by the participants while performing semi-naturalistic tasks with map-based data. We compared two glyph visualization techniques, the shape-based Polyline, and the colour-based Mondrian, to understand their influence on scanning behaviours. The control tech-

nique, Polyline by Opach et al. [OPDR18], is a shape-based technique. It expresses multivariate values as zig-zagging lines cutting through squares, resulting in a glyph field consisting of many small line plots. We compared the technique against Mondrian (named after Piet Mondrian's artworks). Unlike Polyline, Mondrian is a colour-based technique, which arranges glyphs in a radial composite to express multivariate data. We designed Mondrian to support the detection of trends and outliers by scanning the glyph field. Additionally, when two or more Mondrian glyphs overlap, the colour of the overlapping area can express composite data (e.g., by multiplying or subtracting overlapping values).

Participants completed semi-naturalistic tasks based on geospatial regression analysis. The participants scanned the glyph fields using the AR+tablet interface displaying real-world map data. They then indicated the statistical information. Our study design is motivated by Whitlock et al. [WWS20] who argue that geospatial analysts are interested in analyzing geospatial in the wild, and create in-situ large-area visualizations of the data at the same time. They also argue that combining mobile devices and AR is helpful for this endeavour. In the study, our participants used the tablet primarily as an input device for panning the AR content, and to reference place names that were not displayed in AR. They used AR for wide-area scanning and exploration.

Polyline elicited more tablet-based panning of the glyph field—indicating that the technique encouraged the participants to look at the individual glyphs more closely. Mondrian tended to elicit more gaze trajectories; the participants often examined the glyphs from afar, bringing them closer when reviewing the smaller overlapping regions of Mondrian glyphs. Despite having to scroll more, NASA-TLX and SUS scores favour the Polyline technique, possibly due to hardware-based colour distortions with Mondrian. This study demonstrates that visual channels like shape, and colour, can be used to manipulate the glyph field navigation behaviours. This has implications for room-sized glyph field navigation. Furthermore, although an AR+tablet interface seemingly follows focus+context (F+C) due to different display resolutions of the OST-HWD and the tablet [BGS01], we often observed the participants ignoring the tablet. Having two different display resolutions is not a sufficient condition for F+C.

## 2. Related Work and Background Information

### 2.1. Immersive Analytics

According to Dwyer et al. [DMI\*18], immersive analytics is an interdisciplinary field extending the knowledge in information visualization, and visual analytics through by incorporating immersion. Immersion can be added through mixed reality technologies, data physicalization, and multisensory presentation—to name a few. Immersive analytics with AR have been incorporated in some geospatial work [ÇGS\*20] – e.g. Tangible Globes by Satriadi et al. [SSE\*22], and FieldView by Whitlock et al. [WWS20]. Our work advances immersive analytics as it involves navigation with large room-sized glyph fields.

### 2.2. Glyph-based Visualization

Glyph-based visualization involves the use of visual markers to convey information [BKC\*13]. According to Borgo et al. [BKC\*13], we can use visual semiotics to manipulate different aspects of glyphs' appearance to convey different information. Furthermore, we can combine multiple glyphs that possess the same location into a composite to convey multivariate information [ROP11]. As Mondrian is a colour-based technique, some discussion on colour perception is necessary. Unfortunately, colour perception studies with OST-HWDs like the HoloLens are few [EKBW20], perhaps due to colour distortion associated with such device type [ILSP21]. Therefore, we must fall back on information visualization literature for colour-based perception. Literature in information visualization suggests that the colourmap should be “sortable” – i.e. a glyph with a lower value should have less luminance than the one with a higher value [CSH20], and to avoid the use of the rainbow colourmap [CSH20, Mor09].

One form of glyph-based visualization incorporates *small multiples*. Small multiples, essentially, are small charts (e.g., bar charts [PAPB20], parallel plots or line charts [OPDR18, OR18], or radar plots [CLGD18]) designed to display specific statistical information [vdEvW13]. Polyline, containing a zig-zagging line cutting through a square, is a shape-based technique—as the line contains some thickness. Fuchs et al. [FIBK17] demonstrate that we can arrange multiple glyphs and their composites in multiple ways: for

example, we can use a grid-like structure, or arrange the glyphs using geospatial information. Glyphs can be superimposed onto a map. Opach et al. [OPDR18] evaluated Polyline using both grid-based and geospatial placement. They found that geospatial placement performed better because the participants could additionally rely on spatial information to make decisions. Grid-based placement eliminates geospatial information. AR allows for more types of placements, such as around the user using a spherical and other non-Cartesian coordinate system [LPED20, DM22]. To the best of our knowledge, our work is the first to analyze trajectories of small multiples on maps displayed on a large AR-based display area.

### 2.3. Virtual AR Display

OST-HWDs are useful for extending the interface of other devices. Despite small field-of-views of their screens [GEA\*17], they are highly mobile [PNB\*21, KC17, WSS20]. MARVIS, by Langer et al [LSBD21], uses AR to add 3D elements to a tablet's workspace. Meanwhile, Pavanatto et al. [PNB\*21] demonstrate how AR can provide additional monitors to the user while allowing them to maintain mobility. Whitlock et al. [WSS20] found that AR display also encourages more navigation. However, the colour perception is less accurate with OST-HWDs [WSS20, ILSP21]. AR displays tend to have inferior display resolutions to other devices like tablets. Therefore, if we use both AR and a tablet as a display device, this may follow the F+C paradigm. Baudisch et al. [BGS01] developed F+C as a way to conserve graphical resources. Two display resolutions are used. A high resolution is used where the user is focusing. A lower one provides contextual information.

### 2.4. Input Methods

AR input methods and AR graphic user interface do not support precise input [SBSPM21, FS91]. In an early work, Feiner & Shamash [FS91] proposed the use of a hybrid interface where, in addition to the AR device, the user has another input device for high-precision tasks. Since then, there have been some advancements in non-hybrid AR input (e.g., gesture, voice command). However, Soares et al. [SBSPM21] still found that these input methods are still inaccurate. Some researchers, like Satriadi et al. [SSE\*22] introduce new input devices like Tangible Globes to enhance input for AR-based geospatial analysis, while others use a mobile device for precise input: for example, FieldView by Whitlock et al. uses a smartphone [WWS20], and STREAM by Hubenschmid et al [HZBR21] uses a tablet.

### 2.5. Multiple Linear Regression and Tasks

The study task involves analyzing multiple linear regression (MLR) models with pre-selected independent variables (IV) and dependent variables (DV). A MLR model essentially involves creating an equation with a DV being a linear combination of IVs. Participants also analyzed pre-fitted models. Based on prior literature [Bra04, Dao17, DK18, Fri82, JSRC\*21, JBPWM06, LBS90], we argue performing MLR involves these criteria:

**Parsimony.** Does a model have too many IVs? If multiple IVs are highly correlated, some should be removed. Otherwise, a fitted

model will not be parsimonious and can have issues like multicollinearity [Dao17].

**Multiplicativity.** Considering if there is any Multiplicativity (a.k.a interaction effect) between the IVs, or if each IV's effect is independent. Not considering multiplicativity can affect a model's accuracy [Bra04, Fri82].

**Correlation.** The IVs must be able to explain the variances of observed DV values. Example effect sizes include  $R^2$  and adjusted  $R^2$  to quantify this [LBS90]. In visualization, if there is a correlation between two or more variables, a change in one variable should also be observable in another. For instance, when the values of one variable are observed to be increasing, the values of another variable could be increasing or decreasing. Correlation may not exist if the values of the first variables do not predict the values of the others.

**Goodness-of-fit.** The values predicted by the equation must fit the observed DV values. We can use likelihood to measure this [JBPWM06].

**Spatial Autocorrelation.** Data may cluster spatially, which may complicate fitting; for instance, overall national statistics can differ at state-level [PAPB20].

Actions such as data cleaning, feature selection, and model assessment help us to understand a model [DK18]. Here, we introduce the terms “pre-fit” task, and “post-fit” task to categorize these actions. A pre-fit task includes actions performed before fitting a model. A post-fit task is for after fitting the model. These tasks form the basis of our study procedure.

### 3. Study Design

#### 3.1. Research Questions

**RQ1: How do the techniques affect the user's scanning behaviours?** Scanning is important for navigating and understanding glyph fields. We explored how each technique affected various types of trajectories possible within an AR+tablet interface. Polyline is good for identifying trends [OPDR18]. Meanwhile, Mondrian, as a colour-based technique, is good for pre-attentive perception [ROP11]. Unlike Polyline, Mondrian can express additional information which can affect scanning behaviours.

**RQ2: What is the cognitive load associated, and the usability of each technique?** Since Polyline is shape-based and Mondrian is colour-based, we expect the user to have different experiences. We hypothesize that the user will have a better experience with Mondrian since Polyline requires the user to be closer in order to comprehend it. We administered self-reported questionnaires like System Usability Scale (SUS) [Bro95], and NASA-TLX [Har06]. Additional interviews supplement the questionnaires.

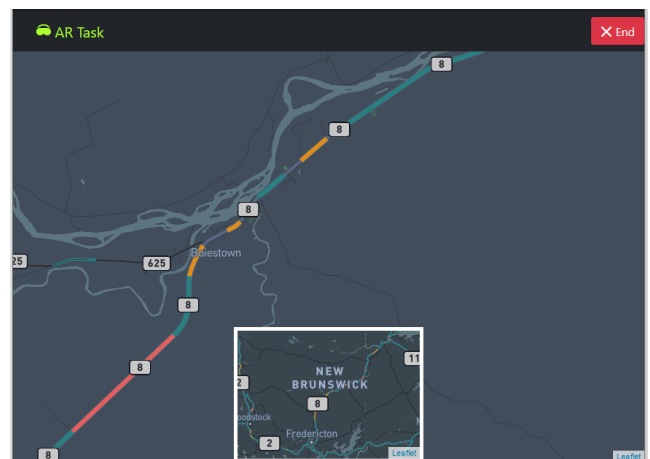
#### 3.2. Participant

We recruited 24 participants (20 males, 4 females) using our institution's mailing lists. All participants were undergraduate and graduate students in computer science, with some familiarity with MLR.

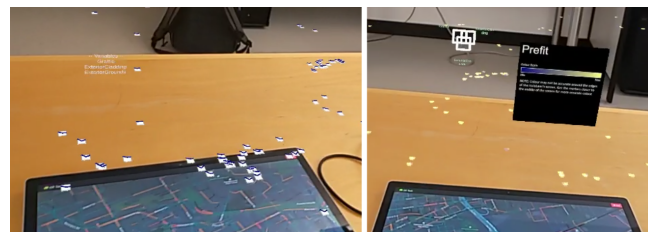
One participant completed secondary education, 15 completed undergraduate studies, and eight possessed graduate degrees. Twenty-two participants had prior experience with mobile AR. Eight had experience with OST-HWDs. Seven had experience with virtual reality. The experience can be overlapping—ie. one participant could have experienced more than one technology.

#### 3.3. Software and Instrument

We used a modified version of Gander for the study. Gander is our AR+tablet interface for geospatial analysis. It reads comma-separated value (CSV) files, and for each data point, generates a glyph. The glyphs' positions are placed based on their geolocations (i.e. latitude and longitude). Gander places these glyphs in AR—creating a room-sized visualization. The appearance of the glyphs changes between the pre-fit and the post-fit stages. Glyph fields' backgrounds are transparent, so the tablet provides additional visual information (e.g. location name, road, etc.) to the user. The user can scan their head to view the glyph field, scroll on the tablet to pan the field, or walk around. The user can examine the glyphs before fitting a MLR model (pre-fit task), and examine the likelihoods after (post-fit task). A “kiosk-mode” version of Gander was used in this study for more experimental control. To assist the user,



**Figure 1:** The tablet interface. The smaller inset is the virtual touchpad, which allows for rapid scrolling of the entire area. Tapping “End” means all sub-tasks have been completed.



**Figure 2:** Modified Gander with glyphs and legends. **LEFT.** P11 with Mondrian. **RIGHT.** P14 with Polyline.

a legend is available in AR; its appearance varies based on the techniques (Fig. 2).

In the study, we used a 15-inch Microsoft Surface Book 3, laid horizontally, displaying a map. Microsoft HoloLens v2 mounted a glyph field on the top. Glyph fields in the study were larger than the tablet which means most glyphs appear beyond the boundaries of the tablet's screen (See Fig. 2). The participants could navigate around the glyph field (i.e. standing up, or walking), or they could use a virtual touchpad on the tablet to pan the glyph field. We did not implement zooming, since it could act as a confounding factor in analyzing scanning behaviours.

### 3.3.1. Glyph Design

We compared two glyph-based visualization techniques: Polyline and Mondrian to represent information during the pre-fit and the post-fit tasks. Each Polyline glyph or a Mondrian glyph composite represents a data point. Each glyph's position is based on the associated data point's geolocation. During the pre-fit task, each represents normalized values. Meanwhile, during the post-fit task, each represents likelihood effect size  $E_L$ .  $E_L = \frac{\ell_{i,j}}{\max \ell}$  where  $\ell_{i,j}$  is a single MLR likelihood for the model  $i$  and the data point  $j$ .  $\max \ell$  is the maximum likelihood for both models.  $E_L$  is bound between zero and one. Our version of  $E_L$  is based on the work by Johnston et al [JBPWM06].

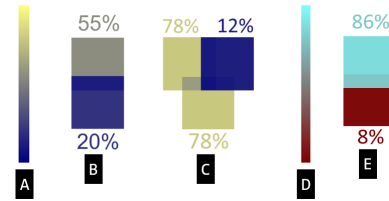
Polyline (Fig. 2-LEFT, Fig. 3) is a small-multiples technique based on the line chart. Since the line is zig-zagging, one can consider this a shape-based technique. In the pre-fit stage with two variables, the line's start point on the left indicates the value of the first variable. Meanwhile, the endpoint indicates the value of the second variable. Each value is normalized based on its variable. For instance, if a variable's minimum is 0 and maximum is 100, the value of 60 is normalized to 0.6. If there are three variables instead of two, the starting point on the left represents the first variable's value, the middle point represents the second variable's value, and the ending point represents the third variable's value. In the post-fit stage, the starting point represents the  $E_L$  of the restricted model at that data point and the ending point represents the  $E_L$  of the full model.



**Figure 3:** The designs of Polyline. **LEFT:** Polyline for the two-variable pre-fit task. The first variable's normalized value is 0. The second one is 0.5. **MIDDLE:** Polyline for the three-variable pre-fit task. The first value is 1. The second one is 0. The third one is 0.67. **RIGHT:** Polyline for post-fit. The first model (left)'s  $E_L$  is 0.25 and so is the second model (right)'s.

Mondrian (Fig. 2-RIGHT, Fig. 4) is a colour-based technique composed of multiple coloured squares. Using Ropinski et al.'s terminology [ROP11], we call each group of coloured squares a composite. The data's geolocation determines the position of the glyph in AR. Overlapping areas indicate multiplicativity during a pre-fit task and likelihood differences in a post-fit task. This is different

from Polyline which never conveys such information. The centre position of the composite is based on the data point's geolocation information. There are other ways to arrange the squares. For instance, the squares could be floating on top of each other. However, we kept the arrangement radial to the technique 2D to allow for a fair comparison against Polyline. 3D arrangements, on the other hand, could introduce a novelty effect to the study. In the pre-fit task, the colour of each square represents the value of a variable at that specific location. The value is normalized to be between zero and one—based on the variable's minimum and maximum. In the post-fit task, each square represents  $E_L$ .



**Figure 4:** Mondrian prototypes. Percentages were not visible to the participants. **A:** Colourmap during the pre-fit task—blue is minimum, and yellow is maximum. **B:** Mondrian during the pre-fit tasks with two variables. The top square represents the first variable and the bottom glyph represents the second. **C:** Mondrian during the pre-fit tasks with three variables. The top-left glyph represents the first variable value, the top-right glyph represents the second one, and the bottom glyph represents the third one. **D:** Colourmap during the post-fit task—red is minimal, and cyan is maximum. **E:** The top represents the  $E_L$  value of the restricted model for that data point. The bottom glyph represents the one for the full model.

### 3.4. Protocol

To obtain semi-naturalistic results, our study tasks involved pre-fit and post-fit tasks with real-world data. There were two maps used in the study: TO which was based on Toronto apartment scoring data [Cit21], and NS which was based on Nova Scotia's lake chemistry data [The21b]. These maps represented different use cases: TO is a smaller map with urban data, and NS is a larger map with natural data. Each participant started with a demographic questionnaire. Then, they read a manual on the tasks and learned about the pre-fit and post-fit tasks. We then verbally quizzed the participants to ensure their understanding of the tasks. If they did not pass, we provided additional information. Then, the participant calibrated the OST-HWD's gaze tracker. Finally, based on their PID, we placed the participants in one of the following groups: G1, G2, G3, and G4. The study had a mixed design; each experienced all techniques and maps, but not all combinations of both.

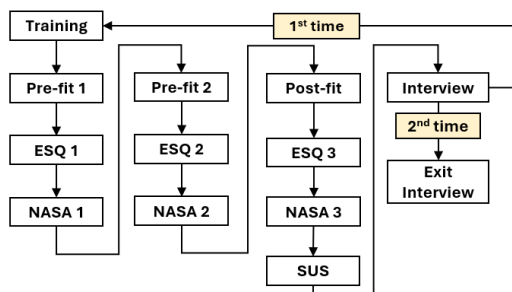
Once well prepared, each participant completed the following steps (also summarized in Fig. 5): **STEP Training.** Trained to perform the pre-fit task. G1 and G4 trained using Polyline while G2 and G3 trained with Mondrian. The training map was [The21a]. **STEP Pre-fit 1.** Using the assigned technique (Polyline for G1 and G4, Mondrian for G2 and G3), performed pre-fit sub-tasks with the assigned map and variables to the group (Table 1-A1 for G1, G2; Table 1-A2 for G3, G4). The sub-tasks were (1) finding minimum



A1: Pre-fit/2 Variables/TO		A2: Pre-fit/2 Variables/NS	
<b>Security:</b>	1-5 score representing security of a building	<b>Calcium:</b>	Calcium level in a lake.
<b>Stairwells:</b>	1-5 score representing the quality of the stairwells.	<b>Chloride:</b>	Chloride level in a lake.
B1: Pre-fit/3 Variables/TO		B2: Pre-fit/3 Variables/NS	
<b>Graffiti:</b>	1-5 score indicating the presence of graffiti on the building. 5 means no graffiti.	<b>Iron:</b>	Iron level in a lake.
<b>Exterior Cladding:</b>	1-5 score indicating the quality of the cladding for the building.	<b>Manganese:</b>	Manganese level in a lake.
<b>Exterior Ground:</b>	1-5 score indicating the quality of the outside area around the apartment.	<b>Potassium:</b>	Potassium level in a lake.
C1: Post-fit/TO		C2: Post-fit/NS	
<b>Model 1:</b>	Score = Graffiti + YearBuilt + Graffiti × YearBuilt	<b>Model 1:</b>	TCU = Iron + Silica + Iron × Silica
<b>Model 2:</b>	Score = Graffiti	<b>Model 2:</b>	TCU = Iron

**Table 1:** The variables/models each participant interacted with in the pre-fit and post-fit tasks. TCU is “True Colour Unit”, a score representing the colour of the water in a lake with particulate matter removed by centrifugation [Hea95].

and maximum values, (2) finding the correlation of the variables to assess parsimony, (3) finding the multiplicative effect between the values, and (4) finding spatial autocorrelation. The participants must verbalize their answers. **STEP ESQ 1.** Using information obtained from the previous step, completed the effect size questionnaire (ESQ), which was a series of multiple-choice questions. See Sec. 3.4.1 for the questions. **STEP NASA 1.** Completed NASA-TLX. **STEP Pre-fit 2.** Repeated *STEP Pre-fit 1* with a different set of variables (Table 1-B1 for G1, G2; Table 1-B2 for G3, G4). **STEP ESQ2.** Repeated *STEP ESQ1*, using information from the previous step. **STEP NASA 2.** Completed NASA-TLX. **STEP Post-fit.** Completed the post-fit sub-tasks for the assigned map and models (Table 1-C1 for G1, G2; Table 1-C2 for G3, G4). The sub-tasks included: (1) assessing the goodness-of-fit of both models, (2) finding the minimum and maximum goodness-of-fit, and (3) finding spatial autocorrelation. Due to the post-fit task being similar to the pre-fit task, no training was provided. **STEP ESQ3.** Completed the effect size questionnaire redesigned for the post-fit task. **STEP NASA 3.** Completed NASA-TLX. **STEP SUS.** Completed SUS. **STEP Interview.** Interviewed for the performance of the technique. **STEP Repeat.** Restarted from *STEP Training* with a different technique; G1, G4 repeated with Mondrian and G2, G3 repeated with Polyline. Also, the map/variables/models changed. G1 and G2 repeated the tasks with NS using the variables/models in Table 1-A2, -B2, -C2. G3 and G4 repeated the task with TO with the variables/the models in Table 1-A1, -B1, -C1. **STEP Exit.** Compared the two techniques in a short exit interview.



**Figure 5:** Summary of the protocol.

We neither encouraged nor discouraged the participants from

walking around the glyph fields. Each session was supposedly 90 minutes long. However, the participants often took up to two hours. Each participant received 25 Canadian dollars. Since the study contained many steps, we developed a web-based portal to serve as a checklist and to help with the administration of the study.

### 3.4.1. ESQ Questions

These questions were based on multiple choice questions often deployed in information visualization studies (e.g., [PAPB20, LKK17]). **For the pre-fit tasks (STEP ESQ1, STEP ESQ2),** the participants completed the following multiple-choice questions: (1) Is there any correlation in the data?, (2) Is there any spatial autocorrelation in the data?, (3) Is there any multiplicative effect in the data?. The choices for the first and second were: *none, weak, moderate, strong*. The choices for the third one were: *very weak, weak, medium, strong, very strong*. **For the post-fit task (STEP ESQ3),** the questions were: (1) What is the goodness-of-fit for Model 1?, (2) Is there any spatial autocorrelation for Model 1’s goodness-of-fit?, (3) What is the goodness-of-fit for Model 2?, and (4) Is there any spatial autocorrelation for Model 2’s goodness-of-fit? The choices for the first and the third questions were: *very weak, weak, medium, strong, very strong*. The choices for the second and fourth questions were: *none, weak, moderate, strong*.

## 4. Analysis and Results

### 4.1. Video Analysis

We conducted two passes on the video recordings, using BORIS [FG16]. In the first pass, we flagged sequences for exclusion in later trajectory analyses (see Section 4.2). In the second pass, we coded sequences of actions. We mostly focused on glyph-field scanning (e.g., looking at the tablet or elsewhere) and navigational behaviours (e.g., walking around, and looking behind). We observed the participants almost always maintained the same forward heading. Those who turned back (P4, P8, P9, P12, P13, P16, P24) only did so temporarily and were aware of the glyphs’ real orientations. Most stayed seated; only P12, P15, P16, P19, P24 stood up at some point. Per participant, about 62% of the scanning and navigational actions were done without looking at the tablet—meaning the AR+tablet interface was not used as a F+C one.

Effects	(a) MSD (Gaze)			(b) MSD (Scroll)			(c) MSD (Position)			(d) Duration		
	$\chi^2$	$p$	$\eta^2$	$\chi^2$	$p$	$\eta^2$	$\chi^2$	$p$	$\eta^2$	$\chi^2$	$p$	$\eta^2$
<b>Technique (T)</b>	4.352	0.037*	0.028	6.619	0.010*	0.062	2.143	0.143	0.014	2.192	0.139	0.009
<b>Map (M)</b>	0.967	0.325	0.006	0.252	0.616	0.003	0.012	0.913	<0.001	0.602	0.438	0.002
<b>Number of Variables or Models (N)</b>	3.789	0.052	0.023	2.547	0.111	0.022	0.000	0.997	<0.001	0.362	0.547	0.082
<b>Post-fit</b>	x	x	x	x	x	x	x	x	x	87.042	<0.001*	0.335
<b>T:M</b>	1.276	0.259	0.034	0.648	0.421	0.008	5.676	0.017*	0.062	0.088	0.766	0.001
<b>T:N</b>	1.289	0.256	0.008	0.033	0.856	<0.001	0.832	0.552	0.002	3.346	0.067	0.013
<b>M:N</b>	0.061	0.805	<0.001	0.189	0.664	0.002	0.832	0.362	0.005	0.005	0.944	<0.001
<b>T:M:N</b>	0.582	0.446	0.004	0.092	0.762	0.001	1.078	0.299	0.007	1.400	0.237	0.005

**Table 2:** Combined ANOVA (Type II, Wald) tables made with log-linked  $\Gamma$  generalized linear models (GLM). The random effect was the participant, all effects had the df of 1. “x” means value unavailable due to fitting errors. Number of Variables or Models means whether the participants were analyzing two or three variables/models— independent of pre-fit or post-fit. Post-fit means whether the AR tasks are pre-fit or post-fit. \* means  $p \leq 0.05$ .

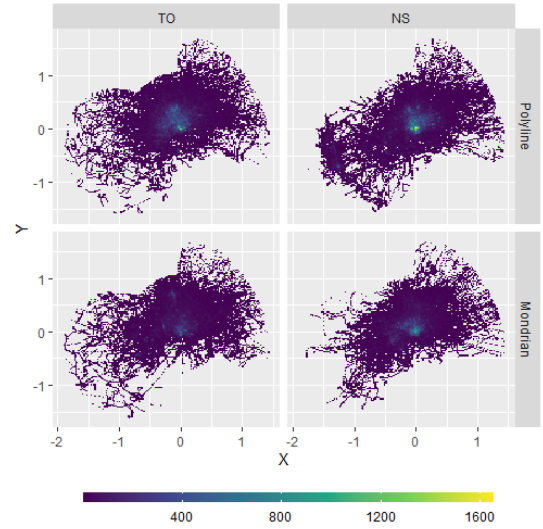
## 4.2. Trajectory Analyses

To answer RQ1, and to understand the participants’ scanning behaviour, we computed three types of mean-squared displacement (MSD): gaze, touchpad scroll (i.e. using the tablet to pan the AR content), and position. We computed MSD for each trajectory using following formula [PFS\*19]: (2D)  $MSD = s^2(X) + s^2(Y)$ , and (3D)  $MSD = s^2(X) + s^2(Y) + s^2(Z)$ , where  $X, Y, Z$  are coordinates of the positions and  $s^2$  is their sample variances. There are other measures than MSD, but they are not usable due to random walk [AVK\*10, MSV18]. Table 2 describes the results of the trajectory tests. Since MSD is not a temporal measure, we performed another ANOVA on the task completion durations. Since all effects have two levels each, a posthoc test is unnecessary. Although the trajectory data are not normally distributed, we report their means and standard deviations because we performed parametric ANOVAs.

### 4.2.1. Gaze Trajectory Analyses

Analyzing gaze data allows us to understand better how the participants looked at the visualization when completing the tasks. Before we can compute MSD, we must pre-process the data. We computed the point of intersection (PoI) between the participants’ gaze rays and the horizontal plane with the tablet at the origin. Since the participants were not always looking at the visualization, some PoIs had extreme positions. We first filtered out any PoI generated when the participants were not looking downward. Then, we used `tclust` [FGEM12] to trim the PoIs with the following parameters:  $k = 3, \alpha = 0.05, iter.max = 100$ . Fig. 6 shows the final results of the conversion process. Finally, we computed 2D MSDs per trial with the trimmed PoIs. Our ANOVA analysis (Table 2-a) shows that only Technique is statistically significant. The effect size is strong with Nakagawa’s  $R_{GLMM}^2$ ’s ( $\psi_1$ ) being:  $R_m^2 = 0.113, R_c^2 = 0.626$ . The result shows the participants’ gazes tended to travel further when using Mondrian—meaning that Mondrian ( $\bar{x}_{MSD} = 0.217m^2, sd = 0.178m^2$ ) tended to encourage more gaze exploration than Polyline ( $\bar{x}_{MSD} = 0.198m^2, sd = 0.174m^2$ ).

Mondrian also has one extra feature that Polyline does not: expression of multiplicative values in the pre-fit tasks and expression of log-likelihood difference in the post-fit tasks through the over-



**Figure 6:** The heatmaps (2D histograms) of combined gaze PoIs grouped by Map and Technique. The Bin width is 0.05. X and Y’s units are metres. Polyline histograms are brighter at the centres due to the participants focusing more around the origin. Mondrian histograms are darker since the gaze trajectories are more spread out.

lapping areas of the glyphs. We tested if the overlapping caused the participants to become more fixated. Let Overlap be the time when the participants were performing the sub-tasks requiring looking at the overlapping area. Before performing an ANOVA, we excluded gaze trajectories with video recordings that were missing (P4+Mondrian+Postfit, P19+Mondrian+Prefit (3 variables)). The ANOVA was Type II (Wald’s test) using a log-linked  $\Gamma$  generalized linear model (GLM) with MSD being the response values. The fixed effects are: Overlap ( $\chi_1^2 = 4.773, p = 0.029, \eta^2 = 0.089$ ), Map ( $\chi_1^2 = 0.859, p = 0.354, \eta^2 = 0.063$ ), and the interaction between the two ( $\chi_1^2 = 0.036, p = 0.849, \eta^2 = 0.001$ ). The random effect is the participants. Nakagawa’s  $R_{GLMM}^2$  ( $\psi_1$ ) are:  $R_m^2 = 0.074, R_c^2 =$

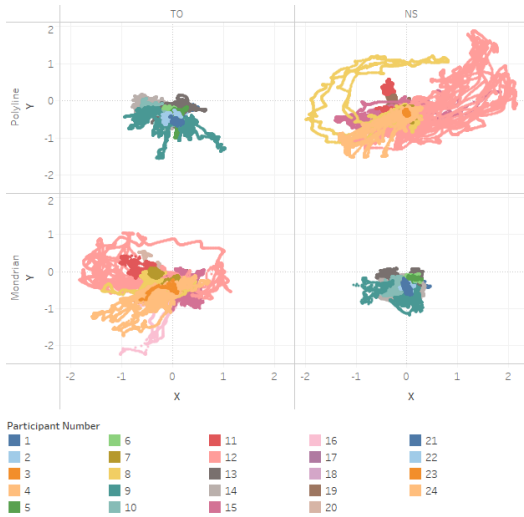
0.726. The test shows that the participants tended to fixate their gaze more when performing sub-tasks looking at the overlapping areas ( $\bar{x}_{Overlap} = 0.198m^2, sd_{Overlap} = 0.151m^2$ ) than when not ( $\bar{x}_{\sim Overlap} = 0.226m^2, sd_{\sim Overlap} = 0.173m^2$ ).

#### 4.2.2. Touchpad Scroll (Panning) Trajectory Analyses

We analyzed how the participants scrolled the touchpad for tablet-based panning (Table 2-b). The video analysis shows there were 12 trajectories where the participants erroneously panned too much and temporarily lost sight of AR content. There are also 25 trajectories where no panning occurred ( $MSD = 0m^2$ ). Both types of trajectories are excluded from the ANOVA test. The test shows that only Technique is statistically significant. Nakagawa's  $R_{GLMM}^2$  ( $\Psi_1$ ) are:  $R_m^2 = 0.093, R_c^2 = 0.324$ . We found that the participants tended to pan the AR content more with Polyline ( $\bar{x}_{MSD} = 0.438m^2, sd = 0.392m^2$ ) than Mondrian ( $\bar{x}_{MSD} = 0.350m^2, sd = 0.404m^2$ ).

To determine if Mondrian's overlap feature affected trajectories, we performed a Type II ANOVA (Wald's test) with a mixed-effect log-linked  $\Gamma$  GLM with MSD as the response value. Only trajectories that met the following criteria were used: (1) no over-panning, and (2) scroll  $MSD > 0$ . The fixed effects are: Overlap ( $\chi_1^2 = 0.002, p = 0.967, \eta^2 < 0.001$ ), and Map ( $\chi_1^2 = 0.015, p = 0.902, \eta^2 = 0.001$ ), and their interaction ( $\chi_1^2 = 0.046, p = 0.830, \eta^2 = 0.001$ ). The participants were the random effect. Nakagawa's  $R_{GLMM}^2$  ( $\Psi_1$ ) are:  $R_m^2 = 0.002, R_c^2 = 0.675$ . The participants did not typically rely on panning to better view the overlapping areas.

#### 4.2.3. Position Trajectory Analysis



**Figure 7:** Combined trajectories of the HoloLens' X and Y positions grouped by Map, Technique, and participant number. X and Y's units are metres.

We computed MSDs of the HoloLen's trajectories. The ANOVA (Table 2-c) shows an antagonistic interaction effect between Technique and Map. Nakagawa's  $R_{GLMM}^2$  ( $\Psi_1$ ) are:

$R_m^2 = 0.165, R_c^2 = 0.654$ . Fig. 7 shows this is due to some participants moving more than others. Due to the partially between-subject of the study, they never completed the tasks with TO+Polyline, and NS+Mondrian. The descriptive statistics were:  $\bar{x}_{TO+Polyline} = 0.016m^2$  ( $sd = 0.024m^2$ ),  $\bar{x}_{NS+Polyline} = 0.150m^2$  ( $sd = 0.278m^2$ ),  $\bar{x}_{TO+Mondrian} = 0.105m^2$  ( $sd = 0.198m^2$ ),  $\bar{x}_{NS+Mondrian} = 0.016m^2$  ( $sd = 0.019m^2$ ).

#### 4.2.4. Trial Duration Analysis

We performed an ANOVA (Table. 2-d) similar to the previous trajectory analyses with an extra factor: Post-fit. Post-fit indicates if the trajectory was pre- or post-fit. The other trajectory ANOVA models do not contain this due to fitting errors. Only Post-fit played a significant role in the trial duration with Nakagawa's  $R_{GLMM}^2$  ( $\Psi_1$ ) effect sizes being:  $R_m^2 = 0.350, R_c^2 = 0.491$ . The mean duration for pre-fit tasks was 290.526 seconds with  $sd = 99.380$  seconds. The mean duration for post-fit trials was 173.932 seconds with  $sd = 76.892$  seconds. This result was indicative of an ordering effect based on tasks as other effects were not statistically significant.

#### 4.3. Self-Reported Measures and Effect Size Questionnaires

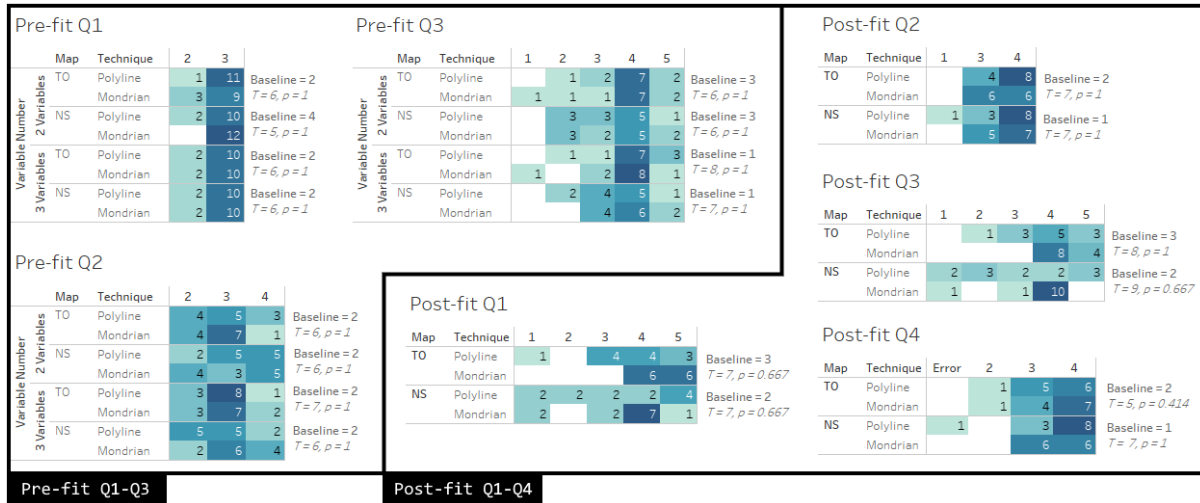
To answer RQ2, we administered NASA-TLX to measure cognitive load. A mixed-effect model parametric ANOVA for NASA-TLX with Technique, Map, Task Types (2-Var, 3-Var, and Post-fit) as the fixed effect and the participants as the random effect; a residual analysis with a QQ-plot determined this type of ANOVA was appropriate. Table 3 describes the results. Only Technique was statistically significant. The descriptive statistics for Techniques were:  $\bar{x}_{Polyline} = 48.793, sd_{Polyline} = 21.472, \bar{x}_{Mondrian} = 52.764, sd_{Mondrian} = 21.582$ . Contrary to our expectation, Mondrian's cognitive load was slightly higher than Polyline's. The interview data indicated that colour distortion may be the cause.

Effects	SS	MSE	df1	df2	F	p	$\eta^2$
T	601.79	601.79	1	109.245	4.739	0.032*	0.039
Map	9.11	9.11	1	109.245	0.072	0.789	0.001
TT	146.96	73.48	2	109.321	0.579	0.562	0.010
TT x Map	99.99	99.99	1	41.267	0.787	0.380	0.007
TT x TT	361.61	180.81	2	109.468	1.424	0.245	0.024
Map x TT	51.85	25.93	2	109.239	0.204	0.816	0.003
T x Map x TT	42.23	21.12	2	109.243	0.166	0.847	0.003

**Table 3:** ANOVA test results for NASA-TLX questionnaire. T = Technique. TT = Task Type (Pre-fit with two variables v Pre-fit three variables v Post-fit). \* denotes  $p \leq 0.05$ .

We administered SUS to measure the techniques' usability. We rescaled the SUS score using Lewis & Sauro's method [LS17] since the 10th SUS question was missed in error. Using ART-ANOVA on SUS with Technique as the fixed effect and the participant as the random effect, we found statistical significance ( $\bar{x}_\Delta = 7.17, F_{1,23} = 6.575, t_{23} = 2.564, p = 0.017, d = 0.741$ ). The median SUS scores were 72.22 for Polyline and 63.888 for Mondrian. The participants found Polyline to be more usable.

The effect size questionnaires (ESQ) were insufficiently sensitive to detect any difference between the two techniques as seen in Fig. 8. We also found that the participants tended to overestimate the effect sizes.



**Figure 8:** The frequency tables of the scores for the pre-fit and post-fit ESQs. 1 = weakest, and 5 = strongest. Baseline numbers and the permutation test statistics for comparing the techniques are provided on the right margin of each table. See [Col13] for more information on the test. Columns with the frequency of zero in all rows have been removed. “Error” column in Post-fit Q4 means one participant did not answer.

#### 4.4. Interview Analysis

To answer RQ2, we analyzed the interview data using the ‘bottom-up’ approach [BC06]. Overall, the pre-fit and post-fit tasks were easy to perform ( $n=4$ ) albeit with an initial steep learning curve ( $n=5$ ). The expanded AR screen was helpful ( $n=3$ ). For the techniques, the participants ( $n=10$ ) thought Mondrian was better for identifying correlation. However, Polyline was easier to understand in general ( $n=13$ ) and felt more precise to use ( $n=2$ ). Fifteen participants indicated that they experienced colour issues with Mondrian. The hardware-based colour distortions produced inconsistent colours based on the viewing angles [ILSP21], forcing the participants to constantly adjust themselves. Some participants ( $n=4$ ) found the divergent colourmaps difficult to understand. P6 believed that more hues (8 to 10) would have been helpful. P12 thought a grayscale colourmap would be better.

#### 5. Discussion and Future Work

Polyline and Mondrian induced different scanning behaviour. The former tends to induce more scrolling on the tablet, while the latter encourages gaze scanning. Therefore, in general, we should use a shape-based technique to encourage closer examination of glyphs, while colour-based techniques are better for encouraging cursory explorations at the pre-attentive level. However, tasks can also affect gaze fixation; we found the participants’ gazes tended to become more fixated on Mondrian glyphs when looking at the overlapping areas. Based on the interviews, we found that the colour-based technique may be less usable due to the colour distortion from an OST-HWD device. Interestingly, some participants suggested the use of a multi-hue rainbow colourmap which is advised against by Moreland [Mor09] and Crameri et al [CSH20]. Future work should explore additional visual channels, and technique-switching in order to optimize glyph field exploration.

Furthermore, we found the participants tended not to rely on the tablet for display. This shows that having a focus+context (F+C) hardware arrangement does not necessarily guarantee a F+C interface. Instead, the interface must incentivize focusing on the tablet. An example of this is blurring the glyphs appearing off the tablet’s boundary, forcing the user to look at the tablet to better view the glyphs. Additional research work should explore other paradigms, such as overview+detail.

The study highlights the need for a new way to measure the effectiveness of glyph visualization techniques in a glyph field. While multiple-choice questions were often deployed in information visualization evaluations (e.g., [PAPB20], [LKK17]), they are too insensitive in the context of tasks involving AR-based glyph field navigation. Furthermore, we may need to move on from discretely categorizing effect size strength. A new and more sensitive method is necessary.

#### 6. Conclusion

We compared two glyph-based visualization techniques for an AR+tablet interface. We found that Polyline induced more tablet-based scroll while Mondrain encouraged visual scanning. This has an implication for designing AR+tablet hybrid user interfaces. We also note that F+C must also be enforced; having a bi-resolution hardware set-up alone is insufficient. This work serves as a template for future multi-device glyph field navigation studies.

#### 7. Acknowledgements

This work was funded by Ericsson–Global Artificial Intelligence Accelerator in Montreal and Mitacs Accelerate grant (IT16687). Joseph Malloch and Muhammad Raza provided feedback that improved this manuscript.



## References

- [AVK\*10] ALMEIDA P. J. A. L., VIEIRA M. V., KAJIN M., FORERO-MEDINA G., CERQUEIRA R.: Indices of movement behaviour: conceptual background, effects of scale and location errors. *Zoologia (Curitiba)* 27, 5 (2010). doi:10.1590/S1984-46702010000500002. 6
- [BC06] BRAUN V., CLARKE V.: Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. URL: <https://www.tandfonline.com/doi/abs/10.1191/1478088706qp0630a>, doi:10.1191/1478088706qp0630a. 8
- [BGS01] BAUDISCH P., GOOD N., STEWART P.: Focus plus context screens: Combining display technology with visualization techniques. *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology* (2001), 31–40. URL: <https://doi.org/10.1145/502348.502354>, doi:10.1145/502348.502354. 2
- [BKC\*13] BORGO R., KEHRER J., CHUNG D. H. S., MAGUIRE E., LARAMEE R. S., HAUSER H., WARD M., CHEN M.: Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications. *Eurographics 2013 - State of the Art Reports* (2013). doi:10.2312/conf/EG2013/stars/039-063. 2
- [Bra04] BRAUMOELLER B. F.: Hypothesis Testing and Multiplicative Interaction Terms. *International Organization* 58, 4 (2004), 807–820. doi:10.1017/S0020818304040251. 2, 3
- [Bro95] BROOKE J.: SUS: A quick and dirty usability scale. *Usability Evaluation In Industry* 189 (Nov. 1995). URL: <https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke>. 3
- [ÇGS\*20] ÇÖLTEKIN A., GRIFFIN A. L., SLINGSBY A., ROBINSON A. C., CHRISTOPHE S., RAUTENBACH V., CHEN M., PETTIT C., KLIPPEL A.: Geospatial Information Visualization and Extended Reality Displays. In *Manual of Digital Earth*, Guo H., Goodchild M. F., Annoni A., (Eds.). Springer Singapore, Singapore, 2020, ch. 7, pp. 229–277. doi:10.1007/978-981-32-9915-3\_7. 2
- [Cit21] CITY OF TORONTO: Apartment Building Registration, 2021. URL: <https://open.toronto.ca/dataset/apartment-building-registration/>. 4
- [CLGD18] CAO N., LIN Y.-R., GOTZ D., DU F.: Z-Glyph: Visualizing outliers in multivariate data. *Information Visualization* 17, 1 (2018), 22–40. doi:10.1177/1473871616686635. 2
- [Col13] COLLINGRIDGE D. S.: A Primer on Quantitized Data Analysis and Permutation Testing. *Journal of Mixed Methods Research* 7, 1 (2013), 81–97. doi:10.1177/1558689812454457. 8
- [CSH20] CRAMERI F., SHEPERD G. E., HERON P. J.: The misuse of colour in science communication. *Nature Communications* 11, 5444 (2020) (Oct. 2020). doi:10.1038/s41467-020-19160-7. 2, 8
- [Dao17] DAOUD J. I.: Multicollinearity and Regression Analysis. *Journal of Physics: Conference Series* 949 (Dec. 2017). doi:10.1088/1742-6596/949/1/012009. 2, 3
- [DK18] DUDLEY J. J., KRISTENSSON P. O.: A Review of User Interface Design for Interactive Machine Learning. *ACM Transactions on Interactive Intelligent Systems* 8, 2 (June 2018). doi:10.1145/3185517. 2, 3
- [DM22] DAEIJAVAD P., MAURER F.: Layouts of 3D Data Visualizations Small Multiples around Users in Immersive Environments. *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (2022), 258–261. doi:10.1109/ISMAR-Adjunct57072.2022.00057. 2
- [DMI\*18] DWYER T., MARRIOTT K., ISENBERG T., KLEIN K., RICHE N., SCHREIBER F., STUERZLINGER W., THOMAS B. H.: Immersive analytics: An introduction. In *Immersive Analytics*, Marriott K., Schreiber F., Dwyer T., Klein K., Riche N. H., Itoh T., Stuerzlinger W., Thomas B. H., (Eds.). Springer International Publishing, 2018, pp. 1–23. doi:10.1007/978-3-030-01388-2\_1. 2
- [EKBW20] ERICKSON A., KIM K., BRUDER G., WELCH G. F.: A Review of Visual Perception Research in Optical See-Through Augmented Reality. *ICAT-EGVE 2020 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments* (2020). doi:10.2312/egve.20201256. 2
- [FG16] FRIARD O., GAMBA M.: BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* 7, 11 (2016), 1325–1330. URL: <https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12584>, doi:10.1111/2041-210X.12584. 5
- [FGEMI12] FRITZ H., GARCÍA-ESCUADERO L. A., MAYO-ISCAR A.: tclust: An R Package for a Trimming Approach to Cluster Analysis. *Journal of Statistical Software* 47, 12 (2012), 1–26. URL: <https://www.jstatsoft.org/index.php/jss/article/view/047i12>, doi:10.18637/jss.v047.i12. 6
- [FIBK17] FUCHS J., ISENBERG P., BEZERIANOS A., KEIM D.: A Systematic Review of Experimental Studies on Data Glyphs. *IEEE Transactions on Visualization and Computer Graphics* 23, 7 (July 2017), 1863–1879. doi:10.1109/TVCG.2016.2549018. 1, 2
- [Fri82] FRIEDRICH R. J.: In Defense of Multiplicative Terms in Multiple Regression Equations. *American Journal of Political Science* 26, 4 (1982), 797–833. URL: <http://www.jstor.org/stable/2110973>. 2, 3
- [FS91] FEINER S., SHAMASH A.: Hybrid User Interfaces: Breeding Virtually Bigger Interfaces for Physically Smaller Computers. *Proceedings of the 4th Annual ACM Symposium on User Interface Software and Technology* (1991), 9–17. doi:10.1145/120782.120783. 1, 2
- [GEA\*17] GRUENEFELD U., ENNENGA D., ALI A. E., HEUTEN W., BOLL S.: EyeSee360: designing a visualization technique for out-of-view objects in head-mounted augmented reality. *Proceedings of the 5th Symposium on Spatial User Interaction - SUI '17* (Oct. 2017), 109–118. doi:10.1145/3131277.3132175. 2
- [Har06] HART S. G.: Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. doi:10.1177/154193120605000909. 3
- [Hea95] HEALTH CANADA: Guidelines for Canadian Drinking Water Quality: Guideline Technical Document – Colour. In *Publications – Healthy living*. The Government of Canada, 1995. URL: <https://www.canada.ca/en/health-canada/services/publications/healthy-living/guidelines-canadian-drinking-water-quality-guideline-technical-document-colour.html>. 5
- [HZBR21] HUBENSCHMID S., ZAGERMANN J., BUTSCHER S., REITERER H.: STREAM: Exploring the Combination of Spatially-Aware Tablets with Augmented Reality Head-Mounted Displays for Immersive Analytics. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021). doi:10.1145/3411764.3445298. 2
- [ILSP21] ITOH Y., LANGLOTZ T., SUTTON J., PLOPSKI A.: Towards indistinguishable augmented reality: A survey on optical see-through head-mounted displays. *ACM Comput. Surv.* 54, 6 (July 2021). doi:10.1145/3453157. 1, 2, 8
- [JBPWM06] JOHNSTON J. E., BERRY K. J., PAUL W., MIELKE J.: Measures of Effect Size for Chi-Squared and Likelihood-Ratio Goodness-of-Fit Tests. *Perceptual and Motor Skills* 103, 2 (2006), 412–414. doi:10.2466/pms.103.2.412-414. 2, 3, 4
- [JSRC\*21] JENKINS-SMITH H., RIPBERGER J., COPELAND G., NOWLIN M., HUGHES T., FISTER A., WEHDE W.: Quantitative Research Methods for Political Science, Public Policy and Public Administration: 4th Edition With Applications in R. University of Oklahoma, Norman, OK, 2021, ch. 13, pp. 127–133. URL: <https://bookdown.org/josiesmith/qrmbook/multiple-regression-and-model-building.html#model-building>. 2

- [KC17] KRESS B. C., CUMMINGS W. J.: 11-1: Invited Paper: Towards the Ultimate Mixed Reality Experience: HoloLens Display Architecture Choices. *SID Symposium Digest of Technical Papers* 48, 1 (2017), 127–131. doi:10.1002/sdtp.11586. 1, 2
- [LBS90] LEWIS-BECK M. S., SKALABAN A.: The R-Squared: Some Straight Talk. *Political Analysis* 2 (1990), 153–171. doi:10.1093/pan/2.1.153. 2, 3
- [LKK17] LEE S., KIM S.-H., KWON B. C.: VLAT: Development of a Visualization Literacy Assessment Test. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 551–560. doi:10.1109/TVCG.2016.2598920. 5, 8
- [LPED20] LIU J., PROUZEAU A., ENS B., DWYER T.: Design and Evaluation of Interactive Small Multiples Data Visualisation in Immersive Spaces. 588–597. doi:10.1109/VR46266.2020.00081. 2
- [LS17] LEWIS J. R., SAURO J.: Can I Leave This One out? The Effect of Dropping an Item from the SUS. *J. Usability Studies* 13, 1 (Nov. 2017), 38–46. URL: <https://uxpajournal.org/dropping-item-sus/>. 7
- [LSBD21] LANGNER R., SATKOWSKI M., BÜSCHEL W., DACHSELT R.: MARVIS: Combining Mobile Devices and Augmented Reality for Visual Data Analysis. *Proceedings of the 2021 ACM Conference on Human Factors in Computing Systems* (May 2021). doi:10.1145/3411764.3445593. 2
- [Mor09] MORELAND K.: Diverging Color Maps for Scientific Visualization. *Advances in Visual Computing* (2009), 92–103. doi:10.1007/978-3-642-10520-3\_9. 2, 8
- [MSV18] MCLEAN D. J., SKOWRON VOLPONI M. A.: trajr: An R package for characterisation of animal trajectories. *Ethology* 124, 6 (2018), 440–448. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/eth.12739>, doi:10.1111/eth.12739. 6
- [OPDR18] OPACH T., POPELKA S., DOLEZALOVA J., RØD J. K.: Star and polyline glyphs in a grid plot and on a map display: which perform better? *Cartography and Geographic Information Science* 45, 5 (2018), 400–419. doi:10.1080/15230406.2017.1364169. 1, 2, 3
- [OR18] OPACH T., RØD J. K.: Augmenting the usability of parallel coordinate plot: The polyline glyphs. *Information Visualization* 17, 2 (2018), 108–127. doi:10.1177/1473871617693041. 2
- [PAPB20] PEÑA-ARAYA V., PIETRIGA E., BEZERIANOS A.: A Comparison of Visualizations for Identifying Correlation over Space and Time. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (2020), 375–385. doi:10.1109/TVCG.2019.2934807. 2, 3, 5, 8
- [PFS\*19] POUPARD M., FERRARI M., SCHLUTER J., MARXER R., GIRAUDET P., BARCHASZ V., GIES V., PAVAN G., GLOTIN H.: Real-time Passive Acoustic 3D Tracking of Deep Diving Cetacean by Small Non-uniform Mobile Surface Antenna. *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2019), 8251–8255. doi:10.1109/ICASSP.2019.8683883. 6
- [PNB\*21] PAVANATTO L., NORTH C., BOWMAN D. A., BADEA C., STOAKLEY R.: Do we still need physical monitors? An evaluation of the usability of AR virtual monitors for productivity work. *2021 IEEE Virtual Reality and 3D User Interfaces (VR)* (2021), 759–767. doi:10.1109/VR50410.2021.00103. 1, 2
- [ROP11] ROPINSKI T., OELTZE S., PREIM B.: Survey of glyph-based visualization techniques for spatial multivariate medical data. *Computers & Graphics* 35, 2 (2011), 392–401. doi:10.1016/j.cag.2011.01.011. 2, 3, 4
- [SBSPM21] SOARES I., B. SOUSA R., PETRY M., MOREIRA A. P.: Accuracy and Repeatability Tests on HoloLens 2 and HTC Vive. *Multi-modal Technologies and Interaction* 5, 8 (2021). doi:10.3390/mti5080047. 1, 2
- [SGHV19] SURALE H. B., GUPTA A., HANCOCK M., VOGEL D.: TabletInVR: Exploring the Design Space for Using a Multi-Touch Tablet in Virtual Reality. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (2019). doi:10.1145/3290605.3300243. 1
- [SSE\*22] SATRIADI K. A., SMILEY J., ENS B., CORDEIL M., CZAUDERNA T., LEE B., YANG Y., DWYER T., JENNY B.: Tangible Globes for Data Visualisation in Augmented Reality. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 505 (2022). doi:10.1145/3491102.3517715. 2
- [The21a] THE GOVERNMENT OF NEW BRUNSWICK: Student enrolment by school by grade by year, as of September 30 / Effectif scolaire par école selon l'année d'études, au 30 septembre, 2021. URL: <https://gnb.socrata.com/Education-Kindergarten-to-Grade-12/Student-enrolment-by-school-by-grade-by-year-as-of/caq2-krss>. 4
- [The21b] THE GOVERNMENT OF NOVA SCOTIA: Nova Scotia Lake Chemistry Data, 2021. URL: <https://data.novascotia.ca/Environment-and-Energy/Nova-Scotia-Lake-Chemistry-Data/vn55-yjyi>. 4
- [vdEvW13] VAN DEN ELZEN S., VAN WIJK J. J.: Small Multiples, Large Singles: A New Approach for Visual Data Exploration. *Computer Graphics Forum* 32, 3pt2 (2013), 191–200. doi:10.1111/cgf.12106. 2
- [WSS20] WHITLOCK M., SMART S., SZAFIR D. A.: Graphical perception for immersive analytics. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (2020), 616–625. doi:10.1109/VR46266.2020.00084. 1, 2
- [WWS20] WHITLOCK M., WU K., SZAFIR D. A.: Designing for mobile and immersive visual analytics in the field. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (2020), 503–513. doi:10.1109/TVCG.2019.2934282. 1, 2