Eurographics Conference on Visualization (EuroVis) 2016
K.-L. Ma, G. Santucci, and J. van Wijk
(Guest Editors)

Volume 35 (2016), Number 3

Faceted Views of Varying Emphasis (FaVVEs): a framework for visualising multi-perspective small multiples

R. Beecham[1], C. Rooney[2], S. Meier [3], J. Dykes[1], A. Slingsby[1], C. Turkay[1], J. Wood[1] and B.L. W. Wong[2]

[1]City University London, UK, [2]Middlesex University, UK, [3]Potsdam University of Applied Sciences, Germany

Figure 1: *Small multiples summarising spatial (red), temporal (blue) and descriptive (green) signatures of several collections of reported road incident data from London. The three perspectives – space, time and description – are superimposed on one another to form space-filling single graphic composites. They indicate, for example, that incidents involving pedal cycles are highly spatially concentrated around central London, although reasonably evenly so, and typically happen during the daytime and mid-week. This is distinct from incidents involving taxis, which have a similar spatial concentration, but typically happen towards the end of the week and in the evening.*

Abstract

Many datasets have multiple perspectives – for example space, time and description – and often analysts are required to study these multiple perspectives concurrently. This concurrent analysis becomes difficult when data are grouped and split into small multiples for comparison. A design challenge is thus to provide representations that enable multiple perspectives, split into small multiples, to be viewed simultaneously in ways that neither clutter nor overload. We present a design framework that allows us to do this. We claim that multi-perspective comparison across small multiples may be possible by superimposing perspectives on one another rather than juxtaposing those perspectives side-by-side. This approach defies conventional wisdom and likely results in visual and informational clutter. For this reason we propose designs at three levels of abstraction for each perspective. By flexibly varying the abstraction level, certain perspectives can be brought into, or out of, focus. We evaluate our framework through laboratory-style user tests. We find that superimposing, rather than juxtaposing, perspective views has little effect on performance of a low-level comparison task. We reflect on the user study and its design to further identify analysis situations for which our framework may be desirable. Although the user study findings were insufficiently discriminating, we believe our framework opens up a new design space for multi-perspective visual analysis.

1. Introduction

Many datasets have multiple perspectives through which they can be considered – for example space, time and description – and often synoptic visual summaries are required that enable each of these perspectives to be analysed at the same time. However, displaying many perspectives simultaneously on a single screen is challenging. The volume of information to be consumed may be overwhelming and visual interference between views may frustrate comparison, leading to high cognitive load [Mun14,JE12,GAW*11,ED07]. This is likely to be especially true when data are faceted, split on some perspective to form useful groupings, and where several data perspectives must be viewed concurrently within each grouping. A design challenge is thus to provide multi-perspective, space-efficient representations that do not clutter visually or confuse cognitively, even when data are faceted into groups or collections for comparison [LMK07].

We present and evaluate a framework that enables such a concurrent overview. We do this by creating visual representations for each perspective and *superimposing* one perspective view on top of the other to form single, space-filling composites, before juxtaposing composites to facilitate visual comparison between collections.

This approach implies a layering of multiple views that do not share the same attribute space. Munzner [Mun14] has delineated a design space for faceted multiple views. Here, she covers the use of small multiples for comparing single perspectives across many collections, or linked perspective views for comparing multiple perspectives on one or two collections (collections being differentiated using colour hue). Instances where *multiple perspectives* are simultaneously compared across *many collections*, by superimposing perspective views and using small multiples to compare across collections, are not discussed. Javed & Elmqvist are perhaps closest to this when they describe *overloaded* views, which are '*like super*[im]*posed views* [but] *overload the space of one visual representation with another visual representation*' [JE12]. Munzner's omission of overloading is for good reason. The main rationale for superimposition is that it allows comparison of data items on the same coordinate space [JE12]. Superimposing entirely distinct views negates this as a possibility.

We hypothesise that where data are faceted into many collections and concurrent analysis of perspectives is important, there may be advantages to such a superimposition. We envisage graphic composites that combine perspective views, which when arranged as small multiples, can be scanned for comparison. These small multiples provide rich, multi-perspective summaries of a collection: important in analysis situations where there are several perspectives that together characterise a collection.

Whilst superimposing perspectives may support concurrent perception of those views, it also likely results in clutter and cognitive load. We try to address this negative side-effect through both our framework and designs. Specifically, we suggest summaries for data perspectives at three levels of abstraction (e.g. Figure 2). As the abstraction level varies, so too does the amount of visual and informational detail within the perspectives. In designing visual representations of these abstraction levels, we make careful decisions about the marks and encodings used so as to minimise the inevitable interaction between views. We call these composite graphics FaVVEs – Faceted Views of Varying Emphasis.

The contributions of this paper are: *(a)* a framework for multi-perspective small multiples; *(b)* a set of designs built upon this framework and applied to spatiotemporal-thematic event data; and *(c)* an evaluation of the proposed framework via a user study.

2. Design prototype and framework for FaVVEs

In this section, we introduce our design prototype and use a discussion of design decisions as a means to elaborate upon our framework for multi-perspective small multiples. Our framework focuses on geo-located event data, and the view combinations reported were designed using a dataset of crime reports in Chicago. When interrogating high volume crime report data, police analysts wish to quickly identify discriminating groups of crimes based on *where*, *when* and *how* those events happen [Nat08, RRF*13]. Our designs thus focus on spatial, temporal and descriptive information, or perspectives, within these crime reports. We do not suggest that they are generalisable to all data analysis situations. However, the designs may be used with similarly structured data; in Figure 1 we apply the same encodings to a dataset of recorded road incidents in London. Moreover, the process through which design decisions were made, specifically the ideas for generating visually and conceptually distinct abstraction levels and rules for minimising interference between superimposed views, may be relevant elsewhere. We reflect on these and the data transformations when outlining our framework. The discussion that follows is organised around the three abstraction levels that form our framework.

2.1. Highest level of abstraction: summaries of central tendency

At the highest level of abstraction, we provide low clutter summaries of the spatial, temporal and descriptive information in crime reports. We do this through the concept of central tendency [UC14] and create encodings with just two marks, communicating the centre and dispersion from that centre in each perspective.

The *spatial perspective* is represented by standard deviational ellipses [Yui11]. Standard deviational ellipses summarise dispersion from the spatial centre of a collection of points across two orthogonal axes [OU02] and give a single dot and ellipse summarising the mean-centre, dispersion and orientation of a point pattern – in this case a set of crimes.

The *temporal perspective* is represented on a polar coordinate system (Figure 2) and we use *circular statistics* [BC06] to represent central tendency. The reference dot identifies the circular mean day (inside ring) and circular mean hour (outside ring) and the reference line is a measure of circular dispersion around the mean [BC06]. The same approach is used to summarise the distribution along continuous, rather than cyclic, time: the dot represents the circular mean timestamp and the reference line is used to represent the interquartile range.

The *descriptive perspective* appears in the margins: the left margin summarises crime type and the right margin location type. The modal value is identified by an abstraction – its first letter – and

Figure 2: *Left. Designs for each level of abstraction: high abstraction, measures of central tendency; medium abstraction, aggregation into bins; low abstraction, maximum detail necessary. Right. Possible abstraction-combinations: occlusion and interference between views is probable where all perspectives are displayed at low abstraction. Where all perspectives are at high abstraction, the graphics are perhaps most obviously 'single' composites.*

a line measuring the relative entropy [AHZ*14] is used to represent dispersion across categories. With a low entropy value, and thus a shorter line, there is little dispersion across categories and the modal value reasonably accurately describes the crime type or location type for that collection.

2.1.1. Design justification

Our approach and design is consistent with Elmqvist & Fekete's [EF10] guidelines for visual summary. They suggest using summaries that give a sense of the underlying data without exposing that detail and argue for simplicity, or rather parsimonious design, in the visual appearance of summaries.

Central tendency is perhaps the most obvious means of summarising a data perspective and each of our views use a limited number of marks to summarise their underlying distributions. An obvious problem with central tendency, however, is that its success is heavily contingent on the nature of the data it seeks to represent. Data that are bi- or multi-modal are not represented well by central tendency. Notwithstanding these concerns, we argue that central tendency represents a level of abstraction that might be desirable and that a more nuanced structure is revealed at the lower levels of abstraction.

Careful decisions were made about visual encoding at this highest level of abstraction. Since these views will be superimposed, or *overloaded*, '*visual clutter* [and] *visual design dependencies between components* [may be] *significant*' [JE12]. We use colour hue to distinguish between perspectives, selecting colours from the Brewer qualitative palette [HB03], and try to rationalise the detail of any marks used.

2.2. Medium abstraction level: data aggregation and binning

At the medium level, we move away from statistical summaries and expose more detail by aggregating or binning data.

In the *spatial perspective*, we create a regular grid, count local densities at locations across the grid and represent those densities using area (with squares positioned at the grid centre). For the *temporal perspective*, we persist with the polar view, but use length to show the relative number of crimes occurring by day of week and hour of day. Below that, we bin continuous time into days and use a histogram to summarise frequencies across those days. The *descriptive perspective* remains contained within the right and left margins, but we show counts across an aggregate, crime super-type and location super-description, again using length to encode quantity. Each category is also now reported using a three letter shorthand.

2.2.1. Design justification

Aggregation is a common technique when treating spatial and temporal data [AA06] and aggregation according to cyclic time – daily and hourly frequencies – is particularly common in crime analysis [BC06]. The aggregation applied to the description view is clearly constrained by the structured categories that are available. However, switching to the medium level of aggregation in each of these views exposes detail that was not captured by the summaries of central tendency. For example, discriminating spatial and temporal patterns may be identified at this medium abstraction level that are hidden at the highest level of abstraction.

The decision to use length for communicating quantitative value can be justified with recourse to graphical perception theory [CM84]. Since cyclic patterns are depicted in the temporal summaries, the use of a polar coordinate system to represent temporal data – a 24 hour clock on the outside ring and a weekday clock on the inside ring – is plausible. We provide further justification for this decision when discussing the superimposition of views (Section 2.4). An important design constraint at the high and medium levels of abstraction is that we avoid varying colour lightness or saturation. We simply use area and length to communicate quantitative value, thus freeing up colour value for the most detailed abstraction level. We argue that this design consistency helps reinforce and make distinguishable the different abstraction levels. expanded this section

2.3. Low abstraction level: maximum details necessary

At the lowest level of abstraction, we attempt to expose as much detail as is desirable or necessary. For the *spatial perspective*, we use kernel-density-estimation (KDE) to create a continuous surface of spatial densities [OU02]. We also include an outline map of the region. For the *temporal perspective*, we expose more detailed information about the temporal distribution of crimes by creating a two-dimensional representation (a calendar view) and showing the volume of crimes by both hour of day and day of week. Below that, in the view of absolute time, we aggregate now to a more precise temporal resolution – hour of day – and smooth these hourly counts by again using kernel-density-estimation (KDE). In the *descriptive perspective* we reveal the most detailed sub-type description. We also consider sub-location descriptions and embed these inside the bars to form spine plots [Hum07].

2.3.1. Design justification

The title for this abstraction level, 'maximum details necessary', is somewhat nebulous. What is considered *desirable* or *necessary* will depend on analysis context. In this paper, decisions around the detail exposed in these views were partly arrived at by considering the crime analysis domain and particularly the process through which Crime Pattern Analysis [Nat08] is performed.

For the spatial perspective, KDE is a technique frequently used by visually-inclined spatial analysts and can be used to summarise point patterns with varying levels of precision [OU02]. The technique is also commonly deployed in crime analysis [BCH07]. The calendar view used in the temporal perspective provides a form of contingency table that is perhaps highly recognisable. A slight inconsistency here is that, rather than simply exposing more precision around the underlying distribution across days of week and hour-of-day separately, this is a bi-variate representation. Additionally, an inevitable consequence of using colour value to encode quantities is that perception of these quantities is affected where views intersect each another. Finally, on the descriptive view, spine plots are very similar to strip treemaps [BSW02]. The use of height (to represent absolute numbers within a crime and location super-type) as well as width (to represent number within crime and location sub-type) helps with comparing proportional differences across parent categories.

2.4. Superimposition of views

Each of these summaries are combined and together form space-filling graphic composites that we call FaVVEs (Faceted Views of Varying Emphasis).

The decision to superimpose the views means that we need to be especially cautious about how views visually interfere with each other. To manage this interference, we *imply* consistency across perspectives at the different abstraction levels, both conceptually – in the form of data abstraction used – and in terms of design – the encodings that appear in perspectives. This ambition is consistent with Wang *et al.* [WBWK00], who recommend using separate visual representations for separate perspectives.

The polar representation for time obviously 'looks' distinct from the spatial view. Moreover, the fact that a substantial portion of the temporal view is unoccupied means that occlusion of the spatial view is minimised. At the high and medium levels of abstraction, the description view is unobtrusive, occupying only the margins and the horizontal bars in the lowest level of abstraction are very obviously distinct from the temporal and spatial view. As well as selecting very distinct mappings of data to coordinate space – polar coordinates for time, cartesian coordinates for space – we use colour hues selected from a Brewer palette [HB03] to reinforce the distinction between perspective views without unwittingly making any single perspective more visually salient.

We accept, however, that there remain problems with the designed views when they are superimposed and that differently shaped regions might require a different view combination. For example, for the Chicago crime data (see Figure 2), the temporal view at the high and medium levels of abstraction interferes with the northern-most and southern-most regions and perhaps makes more salient the centre of Chicago. This is less of a problem when London is considered, a city with a more square-like shape (see Figure 1). An important observation, then, is that our suggested designs represent an extreme case of our framework – complete superimposition. When applying our suggested framework to other datasets, the visualization designer may wish to make specific design decisions around offsetting views in order to minimise interactions between perspective views based on the idiosyncrasies of the data under investigation.

Despite these judicious decisions around design and layout, the more detail introduced into views the greater potential for clutter and occlusion. This is an inevitable consequence of superimposition, which our proposed framework also attempts to address through "*progressive disclosure*" [WBWK00] of detail through the three levels of abstraction.

2.5. Switching emphasis and abstraction level

The abstraction levels and designs for space, time and description can be thought of as *layers* that persist, but can be selectively attended to when needed [BCS11]. For example, after initial comparison across many small multiples, with abstraction levels set to high, focus may shift to a smaller subsets of collections and on a particular perspective: one may expose greater detail on space, but with some abstracted information on time and description. That this switching of emphasis is flexible and fluid is important to our framework. We envisage situations in which analysts quickly move between different abstraction levels in order to build rich overviews across perspectives.

Some consideration was given to the manner by which these different abstraction levels are introduced, or '*made attendable*' [BCS11]. When switching emphasis between abstraction levels we recommend smooth, animated transitions (as Bartram et al. [BCS11] prescribe). We argue that this technique helps reinforce links between the different abstraction levels, bringing to attention alternative overviews that are exposed as the abstraction level changes.

We also suggest varying the emphasis given to different perspectives within a visualization by manipulating the transparency of those perspectives. For example, through the analysis process a certain perspective, say space, appears to be more relevant and discriminating than time or description. This perspective may be analysed in more detail by introducing summaries at the mid- or lower-levels of abstraction. It may also be desirable to emphasise the spatial perspective independent of the abstraction level. For example, the spatial view might be made more *visually* salient by increasing the transparency values applied to other perspectives.

3. Evaluating FaVVEs

Although informed decisions were made around prototype designs, our framework remains speculative: we could find no literature supporting the superimposition of perspective views rather than their juxtaposition, nor for the use of varying abstraction levels. We therefore conducted a user-study to evaluate our framework and designs. The main assumption that we investigated was **whether superimposed views better support concurrent analysis than juxtaposed views in cases where data are faceted into many collections**. Tests were conducted with participants with some data analysis background. Participants were given a repeated set of analysis tasks involving small multiple representations: one with perspective views superimposed as single composites (FaVVEs), the other using the same encoding but with views juxtaposed. The abstraction levels and underlying data were varied between tests. The survey software code and instructions for running the survey is available from: http://www.gicentre.net/favves.

3.1. Design and procedure

3.1.1. Analysis task

Our framework and designs assume that FaVVEs will support synoptic tasks. We imagine scenarios in which small multiples of FaVVEs are visually compared to identify structures shared across one or more perspectives or to detect outliers on one or more perspectives. Rather than articulating precisely on what perspective and to what extent individual composites differ, our ambition with FaVVEs is to provide encodings that *suggest* in a single graphic a multi-perspective signature or profile for the distribution of the underlying data.

For the user study, we therefore created a single analysis task that encourages this more initial and cursory analysis. Participants were given 18 graphics (superimposed composites or juxtaposed separated views) arranged as small multiples. They were also presented with three graphics that were distinct from one another in data space; these graphics represented the centres of three distinct groups. Participants were asked to assign small multiples to the groups based on their similarity. No explicit instructions were given on the relative priority of speed over accuracy. However, the importance of speed was implied by a three-minute countdown that appeared in the top right of the test screens (Figure 3). Additionally, in the pre-test instructions, analysts were encouraged to quickly scan the views 'at-a-glance' and were also reminded that in order to complete the tests, inspection of all three perspectives may be necessary.

An alternative task might have been outlier detection: asking respondents to identify amongst a set of small multiples the graphic that is most distinct. This too would require high-level comparison across one or more perspectives. The grouping task was instead selected as it is a key requirement of Crime Pattern Analysis [Nat08], the use case motivating our initial designs. Here, crime analysts must inspect many sets of crimes and collate those that appear similar to one another based on where, when and how those crimes manifest themselves. In reality police analysts do not have pre-defined group centres of crimes in which they are confident. This more contrived analysis situation was introduced in order to reduce the completion time of the survey and allow for a greater number of experimental factors to be tested.

3.1.2. Experimental factors

Participants were assigned to two cohorts: one performed the grouping task using superimposed composites; the other performed the same task with perspective views juxtaposed. Kept constant were the number of small multiples allocated into groups (18, six in each group), the screen space occupied by the small multiples and the number of tests that individuals must perform (13). Maintaining a constant screen space meant that for the juxtaposed case, perspective views were necessarily smaller. If perspective views were made equal size, it would not be possible to fit the same number of small multiples on a single screen between conditions since the juxtaposed graphics would be necessarily larger. Factors that we chose to vary within-subject were the underlying data distributions used to define perspectives (*perspective-change*) and the level of abstraction or detail exposed in the views (*abstraction-combination*).

Groups were defined by introducing a pattern into a single perspective, into two perspectives or into all three perspectives. Where a single perspective was used, this could be a pattern introduced into the spatial, temporal or descriptive perspectives ($s|t|d$) and where two were used, this could be a combination of alterations to the spatial-temporal, temporal-descriptive or descriptive-spatial ($st|td|ds$) perspectives. We selected four ways through which abstraction levels could be varied: high abstraction for all perspectives ($Hs|Ht|Hd$); medium abstraction for all perspectives ($Ms|Mt|Md$); low abstraction for space, medium for time and high for description ($Ls|Mt|Hd$); high abstraction for space, low abstraction for time, high for description ($Hs|Lt|Hd$).

There are many more ways through which *perspective-change* and *abstraction-combination* could be altered. To mitigate against

Figure 3: *Test screens showing juxtaposed (left) and superimposed (right) design equivalents. Group membership is assigned by selecting a button that temporarily appears when graphics are hovered. Once assigned to a group, the graphic's outline changes from grey to the green (group 1), purple (group 2) or blue (group 3). A timer and information on the total number of graphics assigned to groups appears top right. The timer counts down from 180 seconds, after this it continues with a negative sign and in red (right). Notice that the juxtaposed views are necessarily smaller, but graphics are equal in absolute area.*

respondent fatigue and drop-out, we limited the total number of tests participants performed to 13; this had implications on the extent to which we could allow these factors to vary. In our experimental design, we controlled for the fact that patterns in the underlying data may be more easily identified in certain views and at certain abstraction levels than others. For example, data difference expressed only on the high abstraction spatial perspective may be less obvious than when change is expressed only on the temporal perspective. Performance in the grouping exercise may therefore be poorer than in other test cases not due to the superimposition or otherwise of views, but due to the way in which the stimuli are varied given specific view *abstraction-combinations*. Since there are seven ways in which data perspectives can be varied to define groups ($\{s|t|d\}, \{st|td|ds\}, \{std\}$) and four *abstraction-combinations* that we wish to investigate, 28 tests would be required were each of these *perspective-changes* and *abstraction-combinations* to appear for each participant. Assuming that each test takes on average two minutes to complete, this would mean an average completion time of 26 minutes, excluding the training phase. We aimed to keep the entire user study to within 40 minutes and therefore designed for 13 tests. Within these tests participants were exposed to all four *abstraction-combinations* and three categories of *perspective-change* – single perspective, two perspective and three perspective variation. Additionally, when assigning participants to tests, we sampled systematically through the different sub-levels of *perspective-change* ($\{s|t|d\}, \{st|td|ds\}, \{std\}$), ensuring that there was an even number of tests within any configuration of *perspective-change* and *abstraction-combination*.

One distinction between our experiment conditions and those of a 'real' data analysis is that one would typically expect small multiples to be ordered meaningfully on a screen. Such an ordering may result in spatial autocorrelation in perspective values and certain outlier structures thus more easy to detect. To evaluate the effect of such an ordering on performance, we included a thirteenth test, which contained the same configuration of *perspective-change* and *abstraction-combination* as the first test, but with the small multiples ordered according to group membership.

3.1.3. Training, recorded data and participants

Before performing the experiment tasks, participants received a short visual and textual explanation of each perspective view. Afterwards, a series of nine views were presented and participants had to solve several multiple choice questions. To support learning, the correct answer was displayed after each question.

Responses to both the training and the formal tests were logged, as well as other observational data, such as the completion time for each test and participants' interactions (key presses and mouse movements). Additionally, eye tracking was recorded on four participants' tests. Of the 32 completed tests, 27 took place in a lab setting and with a researcher present, five took place at home. The lab tests were conducted on 22" 1080p screens. It was not possible to control the test environment for the five tests that were taken at home. Four of the 'at home' participants received the juxtaposed views, one the superimposed views. Analysis of performance data suggests no systematic difference between participants taking the test at home versus in-lab. Nineteen of the lab test participants were a cohort of students enrolled in a Masters-level Data Science course and a separate group of eight students, two enrolled on a Bachelors Interface Design course and the rest enrolled as PhD students in Computer Science and Geo-Informatics.

3.2. Generating synthetic test data

We wished to generate realistic looking distributions that nevertheless contained three reasonably distinct groups. We explored various ways of arriving at these data and views. One approach might have been not to generate underlying data, but to contrive data distributions by altering the *visual views* directly. Since the views contain varying levels of detail, this approach might have become problematic for the most detailed views. Instead, we created record-level data. The approach was as follows: random data distributions for each perspective were generated and data representing the centres or anchors for the three groups created. Separate data were generated for each test case (configuration of *perspective-change* and *abstraction-combination*). For the categorical data (time and description) these anchors were created synthetically by specifying

Table 1: *Total number of graphics assigned.*

3	*perspective-change* (1-var, 2-vars, 3-vars)	
4	*abstraction-combination*	×
1	additional test – ordered small multiples	+
32	participants	×
416	tests overall	=
18	graphics per test	×
7,488	graphics assigned to groups	=

a target population mean value and variance; for the spatial perspective 'real' data were sampled from a point pattern of crime locations in Chicago, but within a particular spatial extent. Once data for the group anchors were generated, we created the 18 datasets to be assigned to groups by variably introducing additional records drawn from simulated temporal and descriptive data and sampled spatial data. To avoid any learning effects, new data were generated for each unique test-case.

4. Analysis

The experiment resulted in a reasonably large dataset of assigned graphics – both superimposed composites and juxtaposed views (see Table 1). In our data analysis we summarise performance at the test level and study differences under varying test conditions.

4.1. Analysing test performance

We use a single measure to evaluate accuracy: the assignment *success rate*. This is the number of graphic composites correctly assigned to groups per test. In the following discussion we show how this success rate varies between different test conditions and on the main between-subject experiment factor – superimposition (*s*) versus juxtaposition (*j*). We also consider the time in seconds taken to complete each test (Section 4.2).

Across all conditions the difference between the mean success rate for those receiving the juxtaposed views and those receiving the superimposed views was negligible (15.9 *s* vs. 16.0 *j*, *Cohen's d*: <0.1). There was a very small difference in the average time taken to complete the tests between the superimposed and juxtaposed cases (102secs *s*, 107secs *j*, *Cohen's d*: 0.1).

A challenge with using summary statistics on the success rate variable is that a density plot of these scores shows a very strong left skew: e.g. . Such an obvious ceiling effect is common in studies where tasks are easy to complete; a consequence is a lack of discrimination in test results [Sch14]. Rather than treating the success rate as a continuous variable we instead recode it as a binary variable, differentiating between high ($>= 17/18$, 57% of dataset) and low ($< 17/18$) success. We then create a logistic regression model that attempts to predict the *likelihood* of a test resulting in a high as opposed to low success rate, investigating and controlling for a known set of experiment factors (*superimposition* vs. *juxtaposition*, *completion time*, *perspective-change* and *abstraction-combination*).

Starting with the *superimposition* vs. *juxtaposition* factor as a single predictor, we find that superimposing perspective views has no effect on the likelihood of 'high' success. Completion time has an effect and perhaps in the opposite direction to which one might expect: for every 10 second increase in time, the likelihood of the test resulting in high success reduces by 5%. Where groups were defined on two and three perspectives, the likelihood of the test resulting in high success is 1.6 times and 2.4 times greater respectively than compared to tests where groups were defined by varying a single perspective. Another effect was in the *abstraction-combination* variable: compared with tests where all three perspectives were at the high abstraction level ($Hs|Ht|Hd$), the $Hs|Lt|Hd$ combination is twice as likely to result in a high success score.

The effect of *perspective-change* on success rate is logical. Groups are defined by systematically varying perspectives and if a greater number of perspectives are used to define the groups then group membership should be more obviously defined and the tests more easy to complete. It is more difficult to account for the effect of *abstraction-combination*. One argument might be that there is relatively more detail characterising the composites in the $Hs|Lt|Hd$ test case than compared with the $Hs|Ht|Hd$ test case – and thus there was more information available to correctly distinguish the small multiples. However, the $Lt|Ms|Hd$ and $Mt|Ms|Md$ test cases also expose more detail, but our findings do not suggest a credible effect above the $Hs|Ht|Hd$ test case. Another explanation, then, might be that participants struggled with interpreting the circular statistics in the high-abstraction level time view; the $Hs|Ht|Hd$ test case is the only view configuration where this appears.

One means of investigating these effects further is to generate a multiple-variable model, controlling for each variable that we suspect may be discriminating: *superimposition* vs. *juxtaposition*, *completion time*, *perspective-change* and *abstraction-combination*. The contribution of predictor variables to the multiple variable model is summarised in Figure 4. The effect of *perspective-change* and the negative effect of time still exists, though to a slightly lesser extent; the effect of *abstraction-combination* increases slightly (Figure 4).

4.2. Analysing test completion time

After standardising for variation using Cohen's *d* [Coh90], we observe a very small difference in the global mean completion time for the *superimposed* vs. *juxtaposed* group, with the superimposed group performing the test slightly quicker (102secs *s*, 107secs *j*, *Cohen's d*: -0.1). Comparing the average completion time at different stages within the study – the first through to the thirteenth test performed – there was a reduction in completion time after the first two tests. This suggests that there was a slight adjustment where participants developed strategies for completing the task. Analysis of the thirteenth test, where small multiples are ordered according to group membership, shows that this ordering does speed-up completion time. Completion times generally reduce as the test proceeds. Thus, we compare the average completion for the five tests preceding the thirteenth (94*s*) with that of the thirteenth test (80*s*): considering variability in these times, this is a small-to-moderate effect (Cohen's *d.* 0.3).

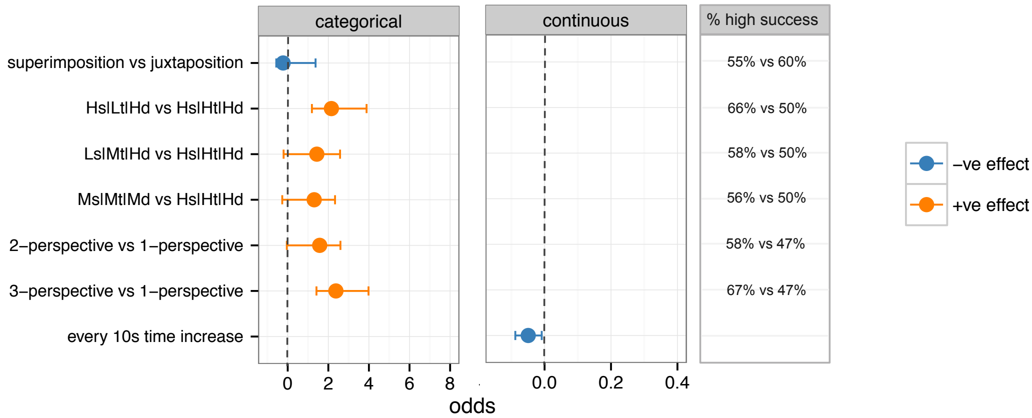

Figure 4: *Multiple variable model output. Outcome: binary success rate. Predictors: s vs. j, abstraction-combination, perspective-change and completion time. Odds-Ratios (exp(b)) and associated 95% CIs are presented. '0' odds represents no effect, thus where the CIs cross '0', there is very little confidence in the estimated coefficient (b).*

4.3. Qualitative insights

That the superimposed tests were completed slightly more quickly, and without affecting performance, *may* be an encouraging finding. The difference might suggest that, unlike the superimposed case where perspectives are combined to form single composites, participants receiving the juxtaposed views had to do some extra work and visual scanning in order to locate perspectives. We additionally performed eye-tracking with four of the 'in-lab' participants; two receiving the superimposed views, two the juxtaposed views. Analysis of the eye-tracking data, given our finding on completion time, was nevertheless inconclusive – it was not clear given the precision of the eye tracking data that participants receiving the juxtaposed in fact performed this additional scanning. Qualitative analysis of the log data also revealed the *strategy* used by participants assigning small multiples to groups: participants almost always moved from left-to-right and top-to-bottom. This more systematic approach is contrary to that envisaged when proposing FaVVEs and constructing the user study and might further explain the lack of discrimination in test results.

5. Discussion

Although we found in our user-study little effect between superimposing rather than juxtaposing perspectives views, we believe our framework still offers potential.

The lack of measurable effect must be discussed in the context of the test environment. That we observe a strong ceiling effect in participants' performance suggests that the design task was insufficiently challenging. This may be due to the synthetically generated data being too clearly defined or the fact that participants had the group centres defined for them in advance – a situation that is unlikely to occur in a real data analysis environment. In reality, analysts may have to interrogate many collections, identify consistent patterns and from these patterns infer links between collections. The decision to include the group centres in our design study was taken deliberately: without their inclusion, the task would be more challenging and time-consuming and variables that we wished to investigate, such as *perspective-change* and *abstraction-combination*, would have been omitted. It should also be noted here that in the final questionnaire section of the test, participants tended towards finding the test challenging and self-reported their own performance as being low. Additionally, varying the number of small multiples that appear on a single screen may have been instructive. It is conceivable that analysts may wish to compare across many more than 18 collections: with many small multiples, and therefore more views across which to scan, there may have been greater differentiation between *superimposition* and *juxtaposition*. Finally, it is highly likely that, in a real scenario, small multiples will be ordered according to a perspective of interest. This ordering may result in visually autocorrelated perspective values due to layout. A hypothesis worth investigating in more detail than in our user-study is whether or not this autocorrelation structure is more easily identified when perspectives are superimposed as single composites.

Whilst it might be possible to investigate these themes with a redesigned user study and a new set of (more challenging) low-level tasks, a more involved evaluation with analysis specialists may be instructive. Specialists might be better placed to answer the second key proposition of our framework, not evaluated in this paper: that designing views at varying levels of abstraction and allowing analysts to selectively bring these perspectives into and out of focus, helps mitigate visual and informational clutter and is useful for concurrent analysis. For example, although our designs can be validated with respect to visual design principles, an open question is whether or not there are situations for which a lower information summary may be useful for analysis. To evaluate this, it would be necessary to consult data analysts who had been exposed to a software prototype for some time and observe whether and how analysts flexibly combine abstraction levels when exploring multi-perspective patterns.

Also worth investigating is how our framework and designs might apply to other data analysis contexts. In mobile applications, FaVVEs may be a means of providing space-efficient, multi-perspective summaries. Or alternatively, in a wider data analysis system, FaVVEs might be used as 'probes' [BDW*08]: positioned at certain geographic, temporal or attribute spaces of interest in order to summarise and monitor multi-perspective activity at those locations.

6. Conclusion

We propose and evaluate a visualization framework that enables analysis of multiple perspectives concurrently, even when data are faceted into small multiples. Our framework suggests that this is possible by superimposing perspective views that do not share the same coordinate space. This superimposition is likely to result in informational and visual clutter. A second argument of our framework is that designing perspective views at differing levels of abstraction, and allowing analysts to flexibly vary these levels of abstraction and detail, may help support concurrent analysis. Our evaluation found that this superimposition, rather than juxtaposition, of perspective views had little effect on a low-level grouping task. The lack of effect might be due to problems of ecological validity in our research design. We reflect on this to suggest real-word situations for which our proposed framework may be most effective and articulate a possible strategy for a more in-depth evaluation. We have identified an analysis scenario that is not accounted for well by existing models of visualization design and, with our framework and accompanying design prototype, have opened up a new design space for multi-perspective visual data analysis.

7. Acknowledgments

This research was supported by the EU under the EC Grant Agreement No. FP7-IP-608142 to Project VALCRI, awarded to Middlesex University and partners.

References

[AA06] ANDRIENKO N., ANDRIENKO G.: *Exploratory Analysis of Spatial and Temporal Data*. Springer-Verlag, Berlin, Germany, 2006. 3

[AHZ*14] ALSAKRAN J., HUANG X., ZHAO Y., YANG J., FAST K.: Using Entropy-Related Measures in Categorical Data Visualization. In *2014 IEEE Pacific Visualization Symposium (PacificVis)* (Mar. 2014), pp. 81–88. 3

[BC06] BRUNSDON C., CORCORAN J.: Using circular statistics to analyse time patterns in crime incidence. *Computers, Environment and Urban Systems 30*, 3 (2006), 300–319. 2, 3

[BCH07] BRUNSDON C., CORCORAN J., HIGGS G.: Visualising space and time in crime patterns: A comparison of methods. *Computers, Environment and Urban Systems 31*, 1 (Jan. 2007), 52–75. 4

[BCS11] BARTRAM L., CHEUNG B., STONE M. C.: The effect of colour and transparency on the perception of overlaid grids. *Visualization and Computer Graphics, IEEE Transactions on 17*, 12 (2011), 1942–1948. 4, 5

[BDW*08] BUTKIEWICZ T., DOU W., WARTELL Z., RIBARSKY W., CHANG R.: Multi-focused geospatial analysis using probes. *Visualization and Computer Graphics, IEEE Transactions on 14*, 6 (2008), 1165–1172. 8

[BSW02] BEDERSON B., SCHNEIDERMAN B., WATTENBERY M.: Ordered and quantum treemaps: Making effective use of 2d space to display hierarchies. *ACM Transactions on Graphics 21* (2002), 833–854. 4

[CM84] CLEVELAND W. S., MCGILL R.: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association 79*, 387 (Sept. 1984), 531–554. 3

[Coh90] COHEN J.: Things I have leaned (so far). *American Psychologist 45*, 12 (1990), 11304–11312. 7

[ED07] ELLIS G., DIX A.: A Taxonomy of Clutter Reduction for Information Visualisation. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (Nov. 2007), 1216–1223. 2

[EF10] ELMQVIST N., FEKETE J.: Hierarchical Aggregation for Information Visualization: Overview, Techniques, and Design Guidelines. *IEEE Transactions on Visualization and Computer Graphics 16*, 3 (May 2010), 439–454. 3

[GAW*11] GLEICHER M., ALBERS D., WALKER R., JUSUFI I., HANSEN C. D., ROBERTS J. C.: Visual comparison for information visualization. *Information Visualization 10*, 4 (Oct. 2011), 289–309. 2

[HB03] HARROWER M., BREWER C.: ColorBrewer.org: an online tool for selecting colour schemes for maps. *The Cartographic Journal 40*, 1 (June 2003), 27–37. 3, 4

[Hum07] HUMMEL J.: Linked bar charts: Analysing categorical data graphically. *Computational Statistics 11* (2007), 23–33. 4

[JE12] JAVED W., ELMQVIST N.: Exploring the design space of composite visualization. In *Pacific Visualization Symposium (PacificVis), 2012 IEEE* (2012), IEEE, pp. 1–8. 2, 3

[LMK07] LAM H., MUNZNER T., KINCAID R.: Overview use in multiple visual information resolution interfaces. *Visualization and Computer Graphics, IEEE Transactions on 13*, 6 (2007), 1278–1285. 2

[Mun14] MUNZNER T.: *Visualization Analysis and Design*. CRC Press, 2014. 2

[Nat08] NATIONAL POLICING IMPROVEMENT AGENCY: Practice advice on Analysis. http://www.acpo.police.uk/documents/crime/2008/200804CRIPAA01.pdf, 2008. Accessed: 2015-03-03. 2, 4, 5

[OU02] O'SULLIVAN D., UNWIN D.: *Geographic Information Analysis*. John Wiley & Sons, New Jersey, USA, 2002. 2, 4

[RRF*13] ROTH R. E., ROSS K. S., FINCH B. G., LUO W., MACEACHREN A. M.: Spatiotemporal crime analysis in US. law enforcement agencies: Current practices and unmet needs. *Government Information Quarterly 30*, 3 (July 2013), 226–240. 2

[Sch14] SCHNALL S.: Clean data: Statistical artifacts wash out replication efforts. *Social Psychology 45*, 2 (2014), 315–317. 7

[UC14] UPTON G., COOK I.: *Oxford Dictionary of Statistics*, 3 ed. Oxford University Press, Oxford, UK, 2014. 2

[WBWK00] WANG BALDONADO M. Q., WOODRUFF A., KUCHINSKY A.: Guidelines for Using Multiple Views in Information Visualization. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (New York, NY, USA, 2000), AVI '00, ACM, pp. 110–119. 4

[Yui11] YUILL R.: The standard deviational ellipse: An updated tool for spatial description. *Geografiska Annaler. Series B, Human Geography 53*, 1 (2011), 28–39. 2