NIH-PA Author Manuscript

A design framework for exploratory geovisualization in epidemiology

Anthony C. Robinson[1]

[1]GeoVISTA Center, Department of Geography, The Pennsylvania State University, PA, U.S.A

Abstract

This paper presents a design framework for geographic visualization based on iterative evaluations of a toolkit designed to support cancer epidemiology. The Exploratory Spatio-Temporal Analysis Toolkit (ESTAT), is intended to support visual exploration through multivariate health data. Its purpose is to provide epidemiologists with the ability to generate new hypotheses or further refine those they may already have. Through an iterative user-centered design process, ESTAT has been evaluated by epidemiologists at the National Cancer Institute (NCI). Results of these evaluations are discussed, and a design framework based on evaluation evidence is presented. The framework provides specific recommendations and considerations for the design and development of a geovisualization toolkit for epidemiology. Its basic structure provides a model for future design and evaluation efforts in information visualization.

Keywords

Geovisualization; user-centered design; epidemiology

Introduction

The promise of information visualization to generate new insights remains compelling. As we are faced with a seemingly endless array of interesting datasets and potential application domains, the pressure to transform individual data elements into accessible synthetic images will only grow greater. This challenge will be met by a combination of new algorithms, rendering techniques, and other technical advances, as well as a set of principles and methods for the effective design and evaluation of information visualization tools. The research described here is focused on this latter challenge.

Since 2003, the Geographic Visualization, Science, Technology, and Applications (GeoVISTA) Center at The Pennsylvania State University has focused evaluation efforts on a long-term project to develop an interactive exploratory toolkit to support cancer epidemiology.[1] The Exploratory Spatio-Temporal Analysis Toolkit (ESTAT), is based on the open-source codeless programming environment GeoVISTA *Studio*. GeoVISTA *Studio* is a Java-based environment designed to provide technically adept users with the ability to create their own customized geographic visualization applications.[2]

ESTAT development has its genesis in earlier work focused on developing modules for commercial GIS software that coordinated multiple views on spatial data.[3] The initial mixture of tools in ESTAT was specified by National Cancer Institute (NCI) based on this

Correspondence: Anthony C. Robinson, GeoVISTA Center, Department of Geography, 302 Walker Building, The Pennsylvania State University, PA, U.S.A. Tel: +1 814 861 8758; Fax: +1 814 863 7943; arobinson@psu.edu.

prior work. The component design and their functional specifications together with the toolkits' application to an epidemiological case study are described in detail elsewhere.[4] This paper focuses on two major portions of the evaluation process undertaken with ESTAT: the results of a verbal protocol study with epidemiologists at NCI, and a design framework that provides both a general structure for future geovisualization design and specific recommendations and considerations for the design of tools to support epidemiology. This research is one component of a larger, multi-investigator, multi-project effort to understand the use of geospatial information visualization methods and to apply this understanding to the design and implementation of methods and tools that are both usable and useful.

ESTAT (Figure 1) features a scatterplot, bivariate map,[5,6] parallel coordinate plot (PCP),[7] and time series graph. Each of these tools is linked to the others so that brushing and selection are instantly coordinated. The ESTAT toolkit is available for download with sample datasets and tutorials at http://www.geovista.psu.edu/ESTAT/.

Epidemiologists stand to benefit from usable and effective information visualization tools. Public health research requires analysts to unravel complex relationships between multiple variables in large datasets, across time and space. ESTAT is designed to support spatio-temporal exploration through multivariate health data. Ideally, users should be able to use ESTAT to develop new hypotheses as well as modify those they may have already created. These design goals were outlined by the NCI in a contract to support development and implementation of ESTAT. It is also a central component of an NCI-funded grant to develop usable and useful analytical methods and tools that integrate visual, statistical, and computational methods.

Developing a workable visualization solution to support exploratory epidemiology is non-trivial for many reasons, including both technological hurdles and difficulties related to determining the best combination of tools for an effective design. User-centered design requires multiple iterations through prototypes, each time acquiring feedback from end users. The goal is to generate a deep understanding of the work that needs to be supported, and tailor tools specifically to those needs, rather than the other way around. This idea often gets lost along the way as we are more eager to create new visualization methods than we are to refine existing techniques for a particular situation. The assumption may be that focusing on iterative evaluation of existing techniques casts revolution aside in favor of the *status quo*. On the contrary, our experiences have shown that users are quick to provide interesting ideas for new methods as they participate in evaluation activities – ideas that have significant relevance toward real-world tasks. ESTAT evaluations have led to insights on multiple aspects of visualization use that have been instrumental in developing the framework presented here.

The potential utility of geovisual methods for exploratory tasks in health research prompts a focus on how these tools can be combined effectively. It is not enough to craft innovative methods – they must be incorporated together in a manner that enables experts in target domains to easily adopt new visual analysis approaches. Our basic framework for the design of information visualization seeks answers to the following questions:

1. What are the features and interactions necessary for geovisualization *tools* that support exploratory health analysis?

2. What are the features and interactions necessary for geovisualization *applications* that support exploratory health analysis?

3. How do epidemiologists use geovisualization in the *analysis* of their data?

> **4.** What *externalities* need to be considered for geovisualization tools to be situated in epidemiological work?

These basic questions emerged during the design of ESTAT, and serve to structure the results of a verbal protocol assessment activity, as well as the set of specific design considerations and recommendations presented here. These elements serve as building blocks for the next generation of geovisual tools for epidemiology and related domains, both in general terms for researchers who wish to evaluate visualization tools, as well as in specific terms for development efforts focused on the domain of epidemiology. The goal of this research is to use the answers gleaned in each of these areas as inputs in the design of future visualization solutions that will more readily fit the needs and constraints of real-world use. Additionally, this research provides an example of a longitudinal design and evaluation effort with an exploratory geovisualization toolkit.

Motivation

Evaluation efforts such as those undertaken with ESTAT can be classified as formative or summative, as defined by Gabbard *et al.*[8] Formative evaluations are focused on user-centered activities and have the aim of iteratively refining designs. Summative evaluations involve direct comparisons between one design and another to draw conclusions about efficiency and effectiveness (usually, to see which option is 'better'). Our work is formative research, as its primary goal was to develop a usable design. It is also inspired in part by Plaisant's[9] recent challenge of 'answering questions you didn't know you had' by allowing users to visually explore and think aloud about what they are trying to accomplish and what they discover during prototypical use.

Researchers in information visualization have been making significant strides toward designing user-centered applications that incorporate geographic data and visualization. Li and North's DataMaps[10] was developed for the U.S. Census Bureau to provide public access to geographic data. It features dynamic query sliders and brushing histograms in coordination with a map to support geographic exploration across multiple attributes. A user test was conducted to determine if sliders or histograms would work better for typical tasks. Li and North conclude that a hybridized brushing histogram which took a few basic features from a query slider offers the most advantages to geographic exploration.

GeoZui3D[11] designed by Ware *et al.* is an environment designed to support 3D exploration of oceanographic data. Presented with the challenge of integrating multiple 3D datasets, Ware *et al.* adopt an approach that emphasizes a sensible interface first (thereby focusing on the user experience), choosing to deal with data issues afterward. Further work[12] has continued to extend GeoZui3D by focusing on 3D navigation through space and time.

A broad range of recent research in geography focuses on the evaluation of geovisualization tools to support a wide range of potential end-users.[13–18] Our work continues this trend toward user-centered tools for geographic visualization. Most empirical evaluations of geographic visualization tools have focused on summative measurements of user performance and preference. Slocum *et al.*[14] diverge from this tradition somewhat by focusing instead on formative work to iteratively design a system for exploring water resource issues. Their experiences indicate that participation by actual end-users throughout the design process was of critical importance when shaping the system. This recommendation matches common practice in user-centered design literature.[8,19,20]

Our work builds upon recent efforts to apply user-centered design methods to exploratory geovisualization tool development. ESTAT evaluation work has integrated users into each step of the design process through several different knowledge elicitation methods over

time. We find that information gathered from a diverse set of evaluation techniques has yielded a rich understanding of the toolkit, where a single experiment using a single technique would have been insufficiently matched to the task of understanding the use of our tools in the complex work of cancer epidemiology. Also, we wish to focus on capturing and building upon ideas from our users, rather than imposing our own.

Evaluating an exploratory toolkit like ESTAT is challenging because its core objectives are to support data exploration and hypothesis generation. Measuring the 'output' of tasks or the 'success' of tool use is not easy. The focus of our evaluation research has been on understanding how epidemiologists take advantage of geovisualization methods and how they expect these methods will augment their daily work. Towards these goals, a variety of knowledge elicitation methods have been combined to iteratively construct our understanding of user's interaction with ESTAT.

The evaluation process

Evaluations have taken place using several usability assessment techniques throughout development (Figure 2). The evaluation process began in October 2003 with a series of rapid assessment activities using card-sorting[19] and verbal protocol analysis[21] techniques. This assessment revealed basic problems with the PCP tool interface. It also clearly indicated a need to shift emphasis to actual end-users rather than the graduate students we had enlisted for this initial evaluation.

This initial evaluation was followed by protocol analysis and focus group activities with 12 users (identified as potential ESTAT end-users by our collaborators at NCI) at the User Centered Informatics Research Laboratory (UCIRL) in Rockville, MD. These users completed a tutorial and set of sample tasks which revealed a serious problem with our data loading/sorting interface. To that point, development focused on the tools and their interactions, rather than the front-end data procedures. We learned from this experience that our users were wishing to change variables often and were generally uncomfortable with the single-panel interface that had been designed to support all data-related tasks. We noted that exploration in epidemiology involves iteration through different subsets of variables, and that data-handling needed to be more dynamic to reflect this tendency. This and many smaller bugs/issues were evaluated to refine ESTAT in preparation for a long-term case study.

The third round of ESTAT evaluations focused on applying the toolkit in collaboration with an epidemiologist from the Penn State Hershey Medical Center over a period of several months. We worked one-on-one with this analyst to examine a problem of his choice, a confirmatory analysis (designed to replicate a recent study he had completed with traditional methods) of colon cancer incidence in Appalachia. His input provided two important insights. First, he identified a set of common general characteristics of epidemiological analysis that we used to implement a set of variable sorting/promotion tools. These tools are based on a strategy that allows analysts to pick variables based on their categorization as outcomes, populations, or indicators – which is the way this epidemiologist had been trained to approach multivariate analysis. Second, he emphasized that this kind of toolkit should be evaluated on the basis of whether or not it could confirm the results of a traditional study. This is important because our design goals to that point had focused on supporting exploration, not confirmation.

Additional results and details from the evaluations leading up to the one reported here are discussed elsewhere.[22,23]

Evaluating ESTAT – individual user task analysis

In December 2004, a task analysis and focus group session was completed at the UCIRL facility. This evaluation was designed to assess the usability of the iteratively refined version of ESTAT. The verbal protocol analysis[21] (VPA) technique was selected for individual task analysis sessions with NCI researchers. VPA has users verbally report their thoughts as they work through one or more tasks. VPA was chosen for this evaluation activity because it can reveal both the suitability of tools for the way users work and the processes involved as they explore and analyze. One decision that must be made when applying VPA is to settle on a response strategy when users get lost or stop talking. Because of the complexity of both the tasks and the toolkit, we decided to offer technical assistance when users were unable to continue. We based this decision on prior evaluation experiences where some users had been stonewalled by a previously unknown bug. When users stopped talking for more than a few seconds, we prompted them to continue.

Following VPA sessions, users were asked to discuss their experiences with ESTAT in a focus group session. Focus groups feature structured group discussion led by a moderator.[24] These discussions are shaped by a set of prompts designed to foster dialog on specific research questions.

Three male and two female participants were scheduled for individual task analysis sessions and a follow up focus group discussion by our project contact at NCI. Each user was an expert health analyst, and within the group they were interested in cancer research on several topics, including: the influence of obesity, tobacco farming, toxicology, and exploring disease burdens.

The protocol analysis portion took place in a room designed for one-on-one evaluation sessions. The UCIRL integrated video system provided recording capability of the test subject's computer screen as well as feeds from several video cameras placed around the room. To facilitate accurate coding of participant actions and transcription of verbalizations, we captured the computer screen and video from a camera focused on participant's faces (Figure 3).

Two weeks prior to the evaluation, participants were asked to download and complete a quick walkthrough of ESTAT using data from the 2004 Presidential Election. This was done to avoid the novelty of the toolkit becoming the focus of each session. At the start of the session, three of the users claimed that they had tried the toolkit as we had requested. The other two users had taken part in the first task analysis sessions at NCI in February, so overall each of the users had interacted at least once with ESTAT on a prior occasion.

Users were asked to complete two tasks in two 40-min sessions. The first task provided participants with a hypothesis that they were asked to either support or refute (Figure 4). Users examined the hypothesis that lung cancer mortality rates are closely correlated with mean annual precipitation (a hypothesis derived from an as yet unexplained relationship among lung cancer, precipitation, and income discussed by Carr[25] and colleagues). County-level data for the lower 48 states were provided for this task.

The second task required each user to pick their own set of variables from a large set of county-level data for Kentucky, Pennsylvania, and West Virginia and to explore these variables with the intention of developing a new hypothesis (Figure 5). The task focused on patterns of ascending and descending colon cancer incidence in this three state region (states that are part of the Appalachian Cancer Network). We chose this task because it reflects emerging research in the epidemiology of colon cancer. Colon cancer in the ascending colon may have a different epidemiology than colon cancer diagnosed in the descending colon.

26,27 This time, users were not given a tentative hypothesis to support or refute. The dataset for task two featured a large and diverse set of outcome variables (mortality and incidence rates for a particular cancer) as well as socioeconomic covariates (such as the percent of persons living under the poverty level).

Each task is detailed online at http://www.personal.psu. edu/acr181/appendices.doc. Following the individual sessions, participants were brought together to discuss their experiences in a focus group session. This was also videotaped, and the session script we used is available at the above web address.

Twelve (two tapes per user and two tapes of focus group discussion) video tapes were transcribed and chunks were coded. The coding categories were imposed on the transcripts based on the four primary research questions outlined in the introduction (Figure 6), where we describe the need to understand *tools*, *applications*, *analysis* methods, and the *externalities* that shape the use of geovisualization tools in epidemiology. These categories are described in further detail as part of the framework that follows this section. The coding categories are also inspired by previous work by Howard and MacEachren[28] who proposed conceptual, operational, and implementation levels for geovisualization interface design. We chose to impose our own pre-determined coding scheme in this case. It is also common to allow schemes to emerge after an initial pass through the data.[29] A pre-determined coding scheme allowed us to search for answers to our particular research questions. We opted not to use an emergent scheme because it would have resulted in coded data that was more closely focused on the peculiarities of ESTAT than on the more general design guidelines for geovisualization we sought.

Comments on design aspects of a single tool were coded into the *Tools* category. Comments directed toward more than one tool at once were coded into the *Application* category. When users described analytical needs or behaviors, these comments were placed into the *Analysis* category. Finally, when users mentioned external factors that shape the way tools are used or perceived, these comments were coded as *Externalities*.

The following sections discuss statements relevant to the four categories we coded and provide a small number of representative quotes to support conclusions. Each major category is further broken down into smaller subsections that address specific areas that emerged from the verbal protocols. At the individual *tool* scale, the scheme is split among each of the tools included in the current ESTAT configuration. At the *application* level, internal and external linkages are described as well as general issues related to composition. Additionally, we coded statements that pertained to *analysis*. Finally, *externalities* are outlined, which in include overarching design and suitability issues.

Individual tools

Scatterplot

During both tasks, participants typically used the scatterplot to drive exploration. This visual method is likely the most common and accessible out of the four available in ESTAT to epidemiologists who generally have educational backgrounds in statistics. Participants appeared comfortable interpreting the scatterplot, and many used it to iterate through numerous variable combinations as part of their analysis. The regression and correlation values provided were valuable, as each participant relied on these values during their analyses to help make choices and modify their hypotheses. One user felt that the statistics provided were insufficiently detailed to support exploration:

P2: I could see a relationship… there's the R-square… it's not giving a statistical significance of the R-square… I can't evaluate that number… because I don't have a confidence interval or something like that.

Bivariate map

The map in ESTAT was frequently used in combination with the scatterplot to iterate through a series of variable pairings. Only one user attempted to use some of the tools included with the map to zoom, pan, and explore using the fisheye lens, for the rest, the map was used only as a dynamically linked overview. Two major issues emerged with respect to the bivariate map. First, users had little or no knowledge of how cartographic classification methods work:

P2: So raw quartiles… I guess I don't know the difference between the meaning of those. And there's no place to go for help to find the definitions of those? I mean, I don't think these are things even as a statistician I… wouldn't know…

Second, identical color schemes were used on the scatterplot and the bivariate map. Thus, the scatterplot serves as a detailed map legend, but this was not obvious to users, several of whom asked questions about the map's bivariate color scheme. These questions only emerged during use of the map. Users were able to interpret patterns in the scatterplot either without considering the color scheme at all, or because the color scheme is superimposed over the values in such a way that it effectively has an integrated legend, thereby permitting users to ignore the 'actual' legend. One of the comments about the bivariate color scheme included:

P3: Labeling the little matrix of colored squares would be helpful… since I happen to know, now it's vaguely coming back to me, that these squares represent something… but you can see the little… the 3×3 square in the corner here is obviously some kind of key, but it's not obvious what it's a key to…

Finally, although ESTAT provides a univariate map option, univariate mapping was problematic for the three users who attempted to modify the variables and classification to work in this way. The univariate map that ESTAT could create at the time was black and white only, and once it had been activated the application usually crashed, making it impossible for users to work with. In hindsight, we suspect that users who are not familiar with advanced cartographic methods would (at least initially) want to work with a univariate map.

Parallel coordinate plot

The PCP required an explanation for three of the five participants. These users had heard of a PCP, but were unable to immediately interpret one. The most common behavior was for users to try the PCP and Time Series graphs after they had explored the scatterplot and map. The relatively low level of familiarity with these visualization methods remains a barrier to adoption, and we should not provide complex and potentially novel tools without easily accessible help functions and training.

PCP exploration generally began with users attempting to reduce the number of lines shown, typically through a small selection across an individual axis or through use of median summary lines (that replaced lines depicting individual counties with lines depicting the median for each state or for each bivariate category into which data are grouped). We fielded more requests for instruction during the use of the PCP than during the use of any other tools, and users expressed frustration that the tooltips provided for each icon in its interface were insufficiently explanatory.

Time series graph

Of the four tools in ESTAT, the time series graph prompted the sharpest complaints from participants. It was built by extending the PCP tool and was aesthetically similar; a strategy that we thought would help learnability and usability in a linked application. For many users, it was too hard to distinguish from the PCP:

> P5: There's the two parts of the parallel coordinate plot… I guess you call the one with time… time series, that one should look as much as possible like a regular graph with an *x* axis and *y* axis…

During the focus group, one user provided a detailed idea of his conception of a time series graph (one that featured a customized interface to replicate results he had seen in an article). Another user augmented this suggestion by pointing out why flexibility should be maintained:

> P1: It depends on what you're looking for… if you're looking for a normal graph versus the clearest way of seeing um… the pattern of the data … there's different elements of a time series… one is the absolute change of the values, we're accustomed to looking at that. Another are coordinated changes, and the scaling choices you make determines what you see… so you should be able to do it both ways…

Discussion

In general, users were able to work through the tasks they were given using the tools in ESTAT. However, there are several key findings on individual tools here. First, users gravitated toward the scatterplot as a starting point in almost every case. Second, bivariate maps are difficult for some users to understand, particularly in terms of deciphering the color scheme. Finally, PCP's, despite their maturity in the information visualization realm, are not widely known or understood by health analysts.

The application

This section outlines the issues raised by task analysis participants that pertained to features at the application level of design (ESTAT as a whole). These include references to internal linkages, external linkages, and the general composition of a toolkit for exploratory spatial-temporal data analysis.

Internal linkages

The coordinated linked brushing that ESTAT features was frequently mentioned as one of its strengths:

> P4: But what I am noticing now is that, as you move from one line to the next I see it moving down here in the map. So that's very nice… so does that happen here… oh yeah! I like this!

While dynamic linking in ESTAT prompted positive comments, participants expected that the tools should be more *fundamentally* linked to one another. Linked-brushing seems to communicate to novice users that the same things are being shown in each view, when in reality they may not be (due to the flexibility of the coordination mechanisms used in ESTAT). Evaluation results suggest that variables in the map and scatterplot should be linked by default – as tested, the default supported greater flexibility (allowing a different pair of variables in the scatterplot and map). Users found it cumbersome to constantly double check to make sure that variable pairings matched between the map and scatterplot (and elsewhere):

> P2: And is there any reason when you change this one 〈mousing over map〉 … this other one 〈mouses over the scatterplot〉 shouldn't be automatically the same?

This theme extends to analysis of time series data. Users found it peculiar that time series data was separated from the primary data, and that they had to search in each set to coordinate their variable choices for exploration:

> P3: It sorta seems like almost anything you select in the time series, you might want it automatically added to the other one, because you're likely to want to look at some of those in relation…

Additionally, selection behaviors in ESTAT were problematic. Most of the participants did not understand immediately why selections (e.g. of a subset of counties) were always maintained, even after the variables displayed had been changed. This functionality may make sense to a geographer, whose focus will be on places and how different variables interrelate for specific places, but it was not intuitive for ESTAT end-users.

Finally, users commonly requested access to full variable descriptions. At the time ESTAT featured only a partial internal linkage that provided metadata. In the data loader and PCP, users could roll over variable names to obtain a description, while this same information was not accessible through the map or scatterplot.

External linkages

When asked if ESTAT should connect to a full-featured statistics package, participants agreed that this was not really necessary as long as the data going into ESTAT was part of a flat file that could be easily imported into statistical software. One user stated:

> P4: … the power of ESTAT to me, I think, is the exploratory end and the linked data… I come out of it with is a better understanding of the interrelationships so that I could then sit down and build some kind of predictive model or something in SAS or some you know, other tool, based on just what I learned and almost the… the output is really just sort of a list of potential useful correlated variables. And their spatial relationships.

Two additional external linkages that were requested included the capability to make screen captures, and methods for sharing projects with other colleagues. At the time of evaluation, ESTAT did not support screen captures (aside from the normal Windows Print Screen function). Similarly, no method for project sharing among users had been implemented. During the focus group, a discussion regarding the need to share visualization results revealed the importance of external linkages to common distribution mechanisms. Users requested export formats that support vector graphics and a mechanism to import what they find in ESTAT into PowerPoint and Word files.

Composition

A recurrent theme throughout this evaluation centered on the general tendency in ESTAT for flexibility in features to take precedence over simple interface design. One user described how simple tasks are made difficult by the overwhelming array of controls provided to modify the time series graph:

> P1: Usually with the rate thing, the common thing I wanna do is have them all scale the same… so you should make that easy to do. Because what I'm sort of feeling a little bit in general with this software is that it's giving extreme amount of flexibility but it's making it hard to do the things I commonly want to do and so… it would be good to sort of make it easy to do the things that most people would want to do…

Also, clicking actions should yield useful results – users expect right clicking to provide a context menu with a set of options. Double clicking should do something commonly needed, like hold an observation or describe it more completely. In the evaluated version of ESTAT, double clicking did nothing on the time series and PCP but it would clone the map and scatterplot. These clones were not interactive, rendering them confusing and unusable – a bug that has since been remedied. One user had a specific suggestion for double-clicking behavior:

> P3: Maybe what I would be hoping for if I double clicked on that point would be it's whole record or something… and then I could, you know the columns would all be labeled and then I could check some things that shocked me…

Selecting and loading variables was not an issue during this assessment compared to prior evaluations. A new data wizard replaced the single-panel data loader. Each of the users made use of the new sorting and categorizing tools (Figure 7).

Selecting variables was a major portion of users' exploration, and they often vocalized their thoughts about possible interesting variable combinations for exploration during this stage. If this is in fact the stage in which an analyst starts to decide 'what would make sense' (as one user put it), then there is a clear need to focus greater attention on making it as interactive, exploratory, and visual as the other tools we provide.

Discussion

The issues raised in this section apply to application design. In ESTAT, the individual tools were built at different times by different programmers and were later assembled into an application. ESTAT takes advantage of the flexibility provided by GeoVISTA *Studio* to mix and match independent components to create custom applications with sophisticated, dynamic, and flexible coordination among components. Asymmetric, distributed development like this can be a problem, however, if the considerations outlined in this section are not heeded during the application design and implementation process. It is crucial to consider the linkages, both internal and external, as well as compositional elements that must come together to form a single package that users do not perceive as a collection of disparate objects.

Analysis using ESTAT

This section summarizes comments made related to analysis using ESTAT. The two tasks selected for users to accomplish were deliberately different. The first prompted them to develop evidence to support or refute a tentative hypothesis, and the second was geared toward providing the opportunity to explore and create their own hypothesis. In general, the first task was handled with greater ease than the second. The dramatic patterns that resulted from the first example were easy to see and explore:

> P3: Anyway, apparently the more it rains the more lung cancer there is… in males. (Figure 8) and… it appears to be true in women too, although not quite as strongly… the r-squared's are… well actually are huge…

The second task was more complicated for several reasons. First, most of the participants did not immediately understand the ascending/descending designations for colon cancer (many incorrectly assuming that ascending and descending referred to upward or downward trends in mortality rates). Second, the dataset for the second task featured many more variables, and the data loading procedure was more complicated because it also featured time series. Finally, the limited geography of the study region (in this case Kentucky, West Virginia, and Pennsylvania) meant that many typical comparisons across racial and economic categories were not possible due to a lack of reliable data.

Individual approaches to task two varied, but it is worth mentioning the character of a few of the attempts because they represent ideas that participants thought they should be able to explore using ESTAT. Two users attempted to explore the relation among ascending and descending colon cancer incidence rates, economic covariates, and access to health care. Another user focused on exploring differences based on race/ethnicity and ascending/descending disease. Yet another had a very specific task in mind, in which he envisioned viewing trends along individual quantiles over time in the time series graph.

Spatial statistic methods were requested by two users who had GIScience expertise. One of these users mentioned the GeoDA spatial analysis toolkit[30] as something he uses for exploratory tasks. Another user offered guidance with respect to the kind of statistics ESTAT should include:

> P3: Yeah, I think the visualization kind of gives you an idea, and then it seems like it'd be a lot of work to add statistical analyses of various kinds, except ones maybe that are really somehow linked to the map.

The user here describes a scenario in which ESTAT emphasizes *spatial* statistics (those that are 'really somehow linked to the map'). Further discussions on this topic revealed a general preference among the group for ESTAT to emphasize spatial statistics and leave traditional statistical analysis to the commercial packages they already use.

Discussion

While users were successful at identifying patterns we had hoped they would explore in the first task, a lack of familiarity with colon cancer research impeded users during the second task. Overall, since users were able to spend most of their time attempting the *tasks* and not struggling with interfaces indicates that the toolkit has progressed substantially since it was first evaluated at NCI.[31] At that time, some aspects of the interface (particularly the data loader) were such an impediment to use that few of the participants were able to accomplish the tasks we had given them in the allotted time.

Externalities

This section outlines the issues that influence tool use but are not generally controllable during the design process. Examples include aspects of organizational culture and domain-specific research traditions. These considerations are external to *development*, but at the same time represent important factors to bear in mind during *design* because they comprise the situation in which tools like ESTAT are utilized.

Situating ESTAT

Some of the focus group questions we asked were designed to foster discussion about how ESTAT (and tools like it) would be situated within the daily work of epidemiologists. We received feedback on a wide range of issues, and here we provide a selection of major themes from that discussion.

To better understand if and where ESTAT can fit within epidemiological work, we asked the group to discuss whether they would accept or resist geovisualization tools like ESTAT. The resulting discussion demonstrates users' desire to have tools that help generate insight – with the caveat that they enable sharing that insight with others:

> P1: I think the test is that we try it on one or two datasets and if it seems to inform or amuse us and give us insight, then we might use it routinely.

P2: Beyond what it does for us, there's the communication of the results that has to be there. We need to be able to take the answer and use it in some way.

P1: But that's that intangible of 'does it give us some insight' that makes us especially enthused about next week's lecture, or webpage illustration, or paper.

P3: I think there is a bias in epidemiology around ecologic data, you know what is it we are looking at, are we sure that it is relevant… what is it we're trying to answer and what are the additional questions that we can gain from the map…

During the same discussion, a user asked the rest of the group what they thought about the spatial focus of ESTAT. He had mentioned his bias toward the aspatial data visualization tools during his task analysis session, stating that, '… the map is something I could get used to using, I guess,' and this comment on the utility of *geographic* visualization resulted in the following exchange:

P5: I think there are maybe two generic kinds of analysis you could do with this software… One is like the kind I was talking about which Gopal did, it's just trends… doesn't use the map at all, just trends by ecologic variables. The other one is more spatial, where you're trying to say what areas of the country have this, and how does these cancer rates relate in a spatial sense… I wonder what is more… useful?

P1: The spatial element is what appeals to me… because the trend, just graphing time series, I can do that in other programs that have a lot more statistical stuff built in… I think what sets this apart for me is the graph-map linkage, and nothing else I have can do that very easily. ArcView is a hard and static program in a way, whereas this… the reason it appeals to me is that it does something that my other statistical thingies don't do.

In this exchange, P1 outlines a specific distinction that makes a tool like ESTAT worthwhile – its dynamism. Traditional mapping software like ArcView is characterized as, '… hard and static,' and P1 provides support for our effort to make geovisual tools exploratory. While there remains work to be done to design ESTAT more effectively for epidemiology, there was a consensus among the small number of users we worked with at NCI that the underlying concepts and tools were well suited for their domain.

Discussion

External factors that were mentioned ranged from practical concerns regarding data creation to more open-ended questions about the changing nature of epidemiological research and the value of exploratory analysis. These issues comprise the situation in which visualization will be used. For this reason, externalities like those mentioned by users in this evaluation activity are important inputs for the design and development of visualization tools. Typically, software evaluation efforts avoid questions about who engineers change, how tools are perceived in the workplace, and other external considerations. These issues are often ignored because they are uncontrollable, but there are gains to be had from evaluating the externalities that impact visualization. For example, knowing that epidemiologists are split about the value of spatial analysis, and that they place high importance on sharing results are important external design inputs.

Evaluating external factors is perhaps the most problematic area of the four highlighted in this work in terms of controlled studies. It is relatively easy to link comments to tools, application issues, and analysis. It is more difficult to separate the wheat from the chaff when users mention external concerns. For this reason we have chosen to highlight the

external factors that came up multiple times (from the task analysis sessions) and were discussed by multiple analysts (from the focus group).

The preceding sections categorize and describe the verbal reports of a small number of users as they have worked with ESTAT to explore cancer data. This study provided valuable insights related to tools, applications, analysis, and external considerations.

Next, we combine the results of all of the ESTAT evaluation activities and crystallize a set of recommendations and considerations for the design of a geovisualization toolkit to supports epidemiology. This framework will inform the future development of visualization toolkits that aim to support health analysis, and it can be taken more generically for other visualization design efforts.

Design framework for an epidemiological geovisualization toolkit

This section presents the condensed end-results of the completed assessments of ESTAT (both those reported here and those reported in Robinson *et al.*31) Arriving at this framework is an interpretive task, and where relevant we note the triangulation from different sources that reinforces a particular recommendation.

The framework is broken down into the four categories used in the previous sections: tool design, application design, geovisual analysis, and externalities. This framework is scale-based from the small scale of individual tools and their features to the large scale of external influences and considerations that an application designed to facilitate exploratory analysis for epidemiology must be situated within. This scheme concisely summarizes the wide array of input gathered during evaluations. The framework is summarized in a graphical hierarchy (Figure 9). This hierarchy is structured to show the ascending scale from elements at the level of individual tools, up to large-scale externalities. At the bottom of the figure, the visualization toolkit is depicted as the total range of issues from each level of scale.

Tool design

At the smallest scale, this framework describes recommendations for features and interactions related to individual tools. Elements in this category represent fundamental features and considerations that are best addressed at the most local scale. The first two sections describe aspects of individual tool use, while the latter five sections describe what has been learned from specific tools in ESTAT.

Interactivity

Most of the positive comments about the individual tools in ESTAT have been related to their fully interactive and dynamic nature. The primary recommendation emerging from this evidence is that geovisualization toolkits should allow high levels of interaction. ESTAT features tools that fit into the *highly interactive* category in Crampton's[32] recent typology of geovisualization interactions. Highly interactive tools are defined as those that include multiple methods of interaction. According to Crampton, the most sophisticated interaction with geovisualizations occurs as analysts attempt to analyze the character of relationships in spatial data. Crampton claims this is best supported by dynamically linked tools, and multiple assessments of ESTAT support this conclusion.

It is important to consider that highly interactive tools, although common in the InfoVis community, are quite novel for health analysts. Specifically, linked-brushing generates excitement among analysts who are used to static representations of their data. During ESTAT evaluations, users were able to manipulate data with simple mouse movements and

actions in order to test variable relationships. This kind of exploratory analysis contrasts sharply with the static statistical techniques that health analysts currently use. As mentioned in the VPA results, linked data selections are problematic when there is no immediately obvious way to 'undo' selections, and when selections are maintained even after variables change. This issue has three potential solutions: we could train users to take advantage of this feature, we could let users control the properties of selection behavior, or we could decide to no longer support persistent selections.

Clonable tools

There are two common approaches to composing a visualization toolkit: one specifies a fixed number of specific tools (as in the current ESTAT application), while the other allows users to reconfigure the number of tools on the fly (a feature of GeoVISTA *Studio*). Based on evaluations of ESTAT, an option in between these two extremes is recommended. A very flexible structure has the potential to impede users with little or no programming experience, but an imposed structure should allow some customization. For epidemiology users this means they could clone tools on-the-fly to create additional maps, scatterplots, etc., to look at multiple geovisual patterns. Maps and scatterplots are especially suitable for this kind of coordinated visual analysis, and it is possible that the PCP and time series graph tools could be stacked on one another (or arranged in trees as presented by Brodbeck and Girardin33) in combinations to examine multiple sections of a large multivariate space at once.

Clonable tools require care in terms of managing screen space. Roberts[34] recently reviewed current strategies for window management, including the use of thumbnail views, elastic window arrangements, and spreadsheet-style organization. One or more of these methods will be implemented in the near future to manage screen space in ESTAT. During evaluations, users frequently mentioned a desire to see the tools in ESTAT on multiple monitors. Designers and developers should consider the fact that applications are often used on a wide variety of computer types, and developing novel methods to manage windows effectively is a necessity in order to compliment exploration and analysis.

Parallel coordinate plots

For most users, the PCP was a new way to visualize data. For that reason it is difficult to determine the utility of a PCP for epidemiology. The main recommendation from ESTAT evaluations is that PCP tools should feature simple interfaces that allow customization after users have become familiar with the technique. During case study work with ESTAT, the epidemiologist we worked with had seen and interacted with PCP's many times before. He was able to use the PCP effectively to compare incidence of colon cancer to a number of different potential covariates. In contrast, during VPA sessions at NCI, most of the users required help to understand how a PCP works. Our inclusion of the PCP tool in ESTAT was driven by a specific request for it from NCI. Therefore an expectation exists at NCI that the PCP may be more widely adopted in the future.

Multivariate analysis using the PCP tools can become visually overwhelming without methods to filter and summarize data. Showing everything can sometimes be quite useful, particularly when examining outliers, but many users were frustrated with the visual 'noise' they experienced while using the PCP to look at observations from all counties in the United States. It was less of a problem in tasks using data from 256 counties in Appalachia. Summary lines to show medians across variables or geographic units are effective tools – provided that some instruction is given to users to describe these features. Users applied these methods during exploratory analysis frequently in place of default views that show all observations at once.

Time series

Epidemiologists want to incorporate time into exploratory analysis using geovisual tools. Accordingly, there is an accompanying need to ensure that temporal graphs are designed to aesthetically mimic common printed time graphs. Interactive time series graphs should take advantage of the common basic features (prominent date labels, a gridded layout, and constant scaling) in static time series graphs. The time series tool we tested was essentially a clone of the PCP tool, including similar icons. We found during evaluations that this similarity was a source of confusion for users.

Designers should carefully consider the implications of adding temporal analysis to geovisualization tools. Users expect that a time series graph will include specialized time series analysis tools. Since the time series graph in ESTAT was derived from the PCP, it featured no methods specifically designed to analyze time. Median summary lines borrowed from the PCP were often used to look at time trends, but users expected additional temporal analysis tools. One user proposed a time series graph that could drive the map, such that users could step through time on the graph and watch the map change accordingly.

Visual data selection

Our evaluation work revealed that there is a need to appropriately address the importance of variable selection in the exploratory process. A visual method for data selection should be part of any geovisualization toolkit that is designed to support multivariate exploration and analysis. While we made major strides with our table-based data loader, we have not yet developed a visually enabled data loader to take its place. This could be done with a simplified version of the correlation matrices recently built for GeoVISTA *Studio*.[35] The variable selection stage of analysis is crucial, and we have not yet reduced the complexity of this task in such a way that it, as well as the visualization tools that follow it, supports exploration. Users responded positively to the idea of using correlation matrices when it was proposed during the focus group discussion.

A major design consideration for a visual data selection tool is the prominent nature of the variable selection process in exploratory tasks. Users suggest that it may be the kind of tool that would persist in the interface, rather than something that is only used prior to the beginning of analysis. In testing, users often returned to the ESTAT data loading procedure during their work to select different variables to view. Since the data loading structure of ESTAT was designed to support use of the module as a step *prior* to exploratory analysis, it was awkward for participants to return to it *during* exploration.

Scatterplots

To support geovisual analysis, scatterplots should provide summary statistics to augment visual patterns with quantitative evidence. Users in our evaluations found scatterplots to be useful and intuitive, but criticized the lack of statistical measures to augment their visual interpretations. Recent versions of ESTAT include correlation and *R*-squared values as well as linear regression lines in the scatterplot. These statistics (and the regression line) are updated when users interact with data (e.g. when a subset of points is selected, the correlation and regression line are reported for the selection), so that comparisons between the entire distribution and subsets of it are possible. Supplementary statistics were essential to the case study collaboration effort, and users in recent evaluations often relied on them.

Many users treated the scatterplot in ESTAT as a legend for the map. Supporting this interaction requires variable pairings to be the same between both components. This is a feature best enabled by default, yet it should be controllable so that advanced users may separate tools more formally and conduct visual analyses that are not necessarily symmetric.

Maps

Two major recommendations emerge from our evaluation results in terms of mapping tools. First, they should support both univariate and bivariate representations. The bivariate map in ESTAT did not provide a reliable and fully functional univariate alternative. This problem emerged during the 'individual' task analysis sessions, as users often tried to begin geographic exploration and analysis by first viewing a univariate map.

The second recommendation is that mapping tools should feature interfaces that provide detailed help information on demand. For those users who did not have prior GIS experience, the lack of help features in the ESTAT map was a barrier to use. Epidemiologists often take a skeptical stance on visual representations until they have developed a reasonable understanding of the underlying data manipulation techniques, and the map was no exception. Users across all evaluations questioned the specifics regarding data classification methods, map projections, and geographic context. For many epidemiologists, maps (if they are used at all) are used primarily to summarize results. Users preferred maps that were easy to create and change like those in ESTAT, as opposed to commercial GIS software.

Application design

The next level of the design framework focuses on issues that affect the design of applications. First, we present recommendations for internal linkages that should exist between tools. Then, we describe the necessary external linkages from an application like ESTAT to other pieces of software. The third section contains guidelines for application composition. These include thoughts on interface design as well as mechanisms for self-directed user education.

Internal linkages

As mentioned previously, linked interactions are important aides in exploration with geovisual tools. Linking should be intuitive and consistent, with sufficient user control over the behavior. The fact that ESTAT did not link variable choices across views automatically was a problem for users in all evaluation settings. There was no time during prior evaluations in which the scatterplot and map needed to show *different* combinations of variables, and users often believed they were either doing something incorrectly, or that the software was not responding to their requests. When we explained the flexibility we had engineered for variable selection across tools, users expressed doubt that that should be the default behavior. Users' analytical work tended to focus on iterating variable combinations in the map and scatterplot, and without linking their variable selection automatically, this task was prone to incorrect interpretation.

A second recommendation is that metadata should be available consistently throughout each of the tools in such a way that users do not realize that the tools in the application were built independently of each other (often a problem with open-source toolkits). Following our initial evaluations at NCI, we added data descriptions to ESTAT's PCP and time series tools as well as to the data loading wizard. During the case study and individual task analysis sessions, users liked this new capability but did not understand why the map and scatterplot would not also provide this information.

External linkages

Two major external linkages should be supported. First, toolkits must provide the ability to capture interesting visualizations to share with colleagues. This opinion was quite strongly held by those who participated in the individual VPA sessions, as they described scenarios

in which they might use geovisualization tools. Capture ability could be initially implemented as static bitmap images of individual windows, but tools should also export vector graphics and lightweight applications that allow limited interactions and some ability to share the history of work that lead to an interesting visual result.

The second important external linkage required is support for exporting a subset of variables from the toolkit to statistical software. This linkage emerged when users were asked whether or not a toolkit like ESTAT should include more statistical analysis tools. Once exploration has resulted in a new or modified hypothesis, users should then be able to easily transition from exploratory visualization to rigorous confirmatory analysis.

Composition

Each of the assessment activities pointed toward a need to conform to a common design philosophy. Geovisualization toolkits should be designed for the simple, frequently used activities and provide users the ability to dig for details. This mirrors common interface design guidelines as described by Shneiderman and Plaisant.[36] It is particularly important that the visualizations themselves are not obscured by the sheer amount of ancillary controls and labels that are visible at the same time. Evidence from multiple evaluation activities suggests that users appreciate flexibility, but desire simplified representations as they start exploratory analysis.

A second recommendation is that tools that have cognates in common use (perhaps in a static form) should have similar aesthetic appearances, while still providing the ability to customize features and modify aspects of the display as a secondary set of options. The time series graph is a specific example of this issue, as users expected to see a time series graph that looked like those they were used to in their daily work. Comments regarding the complexity of ESTAT may not have emerged had it been initially designed with an emphasis on presenting a simple layout that draws upon common interface metaphors.

Thirdly, geovisualization toolkits must provide comprehensive help features, in particular the ability to quickly ask questions of interface features that may be new to users. This could be done by providing clear and understandable rollover text as well as via a query tool much like the 'what's this' feature common in many programs today. Users pointed out that without an expert facilitator, it would be difficult at best to justify the time required to research each tool somewhere else to learn about its usage. Help features should include example applications of the tool that demonstrate the capabilities of exploratory geovisualization, as well as references to published material that describe the tools and methods in greater detail. An expert audience will require guidance toward these sources in order to fully incorporate them into their research.

Analysis using geovisual tools

The following sections describe design considerations related to the types of analyses that occur with geovisual tools, as well as recommendations regarding visually supported spatial analysis methods.

Exploration vs confirmation

Geovisualization toolkits that support epidemiology should provide users with both exploratory and confirmatory capabilities. The case study collaboration focused on how geovisual tools could be used to bolster a traditional analysis with visual confirmation of the results. This task stands in contrast to those we had users attempt at NCI in both February 2004 and December 2004 task analysis sessions. In those instances, we encouraged users to modify an existing hypothesis, or explore data to create and assess a new hypothesis. Our

collaborator in the case study chose to use ESTAT to re-confirm existing research results. The decision to apply ESTAT to replicate traditional results suggests how at least some users in epidemiology might initially approach the use of geovisual tools; they need to be convinced that the tools produce results that complement traditional analysis before trusting those tools to move beyond that type of analysis. This is augmented somewhat by statements made by users who indicated that the test for acceptance will be to show that these types of tools can enable *new* insights about their data. The confirmatory approach is easier for designers and programmers to support. Supporting exploration is a difficult design challenge, and novel evaluation methods are needed to assess that goal.

Spatial analysis

To support spatial analysis, a geovisualization toolkit should focus on providing access to spatial analysis methods that are not available in common statistical software. While there are mixed opinions about the utility of statistical methods in geovisualization, the addition of simple descriptive statistics enabled much of the work carried out in our case study collaboration, and they drove analysis during the VPA sessions. Focus group discussions in December concluded that *spatial* statistics that are not available in other statistical software should be emphasized. This is an immediate analytical advantage a geovisualization tool can provide a health analyst, because they simply do not have access to these methods outside of a full blown GIS, which few are trained to use. Simply adding this functionality will not suffice – these methods are new to most health researchers and require documentation, illustrated examples, and training.

Externalities

The final level of the design framework details external factors outside of the tools, applications, or analytical approaches that have been observed during the course of our evaluations. The two most common external issues are related to databases and user education. A brief section follows this discussion that outlines some of the other external considerations that emerged.

Databases

Creating and maintaining spatial (and spatio-temporal) databases is a major challenge to the widespread adoption of geovisualization tools for health analysis. Geovisualization toolkits should be built with utilities to help users easily create spatial databases and metadata. Our evaluations assumed that users would be handed a database that was ready for immediate use. In reality, health analysts usually create their own datasets, and few of them have sufficient GIS experience such that they could readily spatially join data to boundaries for a tool like ESTAT. While we focus on developing new and better methods for visualizing data, we have yet to seriously tackle the task of enabling the easy creation of multivariate spatial data. Users in our evaluations assumed (correctly) that it would be a difficult task considering the complicated dynamic visualizations they were using.

Web-based data warehouses and lightweight assembly interfaces to access them will likely alleviate this problem, and NCI is working on this kind of tool now for their users who work with ESRI GIS software. Geovisualization toolkits should leverage this development and repackage these existing tools as customized data import interfaces.

User education

User education and training are major hurdles to overcome before these tools become widely adopted. Geovisualization toolkits need on-demand interactive walkthroughs to introduce new users to geovisual exploration and analysis. Interactive visualization tools are quite

different from static methods that health analysts are used to. Virtual 'sticky-notes' as integrated initial help in the interface[37] are a promising method for providing help information. Visualization tools are often overwhelmingly interactive, and typically feature data representations that need to be described before they are used. Documentation in the form of help files can alleviate this to some extent, but interactive walkthroughs are needed to show users how exploration and analysis occurs with geovisualization tools. The PCP is the exemplary case here, as very few of the users across the range of ESTAT evaluations understood what a PCP was or how it functioned. Not only do methods like these need to be disseminated more strategically, their practical use should be more clearly outlined if analysts are expected to incorporate visualization tools into their daily work. One strategy may be to show how the toolkit in question is able to mirror results obtained in a typical analysis, thereby proving its immediate value. Design objectives must be modified in order to do this, however, so that supporting confirmatory analyses is given as much attention as supporting exploration.

Other factors

Design efforts must take into account other external factors that influence the situation in which a toolkit will be used. Users mentioned that the validity of ecological studies was a contentious issue among epidemiologists. Some epidemiologists believe it is impossible to mathematically control for the complexity of the environment, and therefore impossible to conduct meaningful structured studies of spatial phenomena. Another external concern is that users outside of major agencies may have insufficient access to technical support and infrastructure in order to make use of applications like ESTAT. This will be an important consideration as tools like ESTAT are distributed to health analysts in state and local agencies. Finally, users revealed that new tools and methods are often handed down from their superiors. The merits of a toolkit like ESTAT must be obvious to decision makers who influence the adoption of new methods.

Changes to ESTAT

ESTAT has substantially evolved since iterative design and evaluation efforts began. Our development group uses an issue tracking system to monitor and prioritize new features and bugs in ESTAT and to date there have been 83 additions from issues identified during evaluations. Changes range from mundane items such as improved icons, up to major architectural changes involving interactivity between tools, generating basic statistics on-the-fly, and implementing a new interactive legend that allows quick and easy changes to classification and color schemes. In the near future, a redesigned version of ESTAT that features a hierarchical clustering tool will be released for users at NCI, and new evaluation activities are planned for 2008.

A new time series graph interface has been designed that makes it distinct from the PCP (Figure 10) and appears more like commonly used time series graphs. We are rebuilding that data model that underlies the ESTAT tools to support more flexible integration of time series with other data.

To better support the data loading process, we have recently added the capability to click once and skip immediately to the variable selection screen to modify what is visualized. For better drill-down and spreadsheet export features, we are connecting the table browser tool from GeoVISTA *Studio*.

The map has been modified to ensure that switching between univariate and bivariate combinations is easier. Additionally, variable selections are now automatically linked between the scatterplot and map, addressing the issue users had when ESTAT did not couple

variable selection in these two tools. Metadata descriptions now appear when rolling over variable names in all four views.

The next steps for ESTAT evaluation will focus on summative efforts to attempt to more directly measure its contribution to epidemiological work. Recently, Saraiya *et al.*[38] outlined a potential model for empirical measurement of user performance in hypothesis creation tasks with exploratory visualization software. In future summative evaluations of ESTAT, we will consider adopting a similar approach.

Conclusion

Based on iterative evaluation of ESTAT, we have outlined a design framework to guide user-centered design of exploratory geovisualization tools to support epidemiology. This framework, based on work with epidemiologists at NCI, can serve in multiple positions throughout the user-centered design process. During the work analysis stage, a set of considerations and recommendations like those presented here are starting points for the design of an exploratory geovisualization toolkit. Additionally, the general framework structure provides a model for the composition of design and evaluation goals. While it is based on a single case study, the framework presents four general areas of information that tool evaluations should seek.

Further research is necessary in order to evaluate whether or not it is beneficial to use an evaluation approach that from the start aims to acquire information on each of the four major areas of our design framework. The design framework approach could be applied to the design of tools to support public health research to determine whether or not it decreases the number of iterations in the design process before tools are usable and ready for real-world applications. Additional case studies in design for information visualization are needed to expand upon the set of general principles presented here. For example, a similar set of evaluations could be conducted to support a different context of use, such as collaborative geovisualization tools for crisis management, the guidelines that emerge could be compared to identify how well our four major themes apply. Future work might also focus on comparing the differences of design guidelines that come from data that has been coded using emergent schemes versus the pre-determined scheme we have used here.

This research provides guidance for the development of geovisual exploratory tools. This challenge is described by Muntz *et al.*[39] as the development of human–information interaction. Human–information interaction moves beyond the interface of humans and computers, and instead focuses on how people use, acquire, and understand geospatial information itself, not on how people interact with computers to use, acquire, and understand geospatial information.

Acknowledgments

References

1. MacEachren, AM.; Carr, D.; Scott, D. Project highlight: quality graphics for statistical summaries; 6th Annual National Conference on Digital Government Research: Emerging Trends 2005; Atlanta, GA. Marina Del Ray, CA: Digital Government Research Center; 2005. p. 175-176.

2. Takatsuka M, Gahegan M. GeoVISTA Studio: a codeless visual programming environment for geoscientific data analysis and visualization. Computers and Geosciences 2002;28:1131–1144.

3. Edsall, RM.; MacEachren, AM.; Pickle, LW. Case study: design and assessment of an enhanced geographic information system for exploration of multivariate health statistics; IEEE Symposium on Information Visualization 2001; San Diego, CA. Washington, DC: IEEE Computer Society; 2001. p. 159

4. MacEachren AM, Chen J, Robinson AC. Design and application of an exploratory spatio-temporal analysis toolkit. in progress.

5. Olson JM. Spectrally encoded two-variable maps. Annals of the Association of American Geographers 1981;71:259–276.

6. Eyton JR. Complementary-color two-variable maps. Annals of the Association of American Geographers 1984;74:477–490.

7. Inselberg A. The plane with parallel coordinates. The Visual Computer 1985;1:69–91.

8. Gabbard JL, Hix D, Swan JEI. User-centered design and evaluation of virtual environments. IEEE Computer Graphics and Applications 1999;19:51–59.

9. Plaisant, C. The challenge of information visualization evaluation; Conference on Advanced Visual Interfaces; Gallipoli, Italy. New York, NY: ACM Press; 2004. p. 109-116.

10. Li, Q.; North, C. Empirical comparison of dynamic query sliders and brushing histograms; IEEE Information Visualization 2003; Seattle, WA. Washington, DC: IEEE Computer Society; 2003. p. 147-153.

11. Ware, C.; Plumlee, M.; Arsenault, R.; Mayer, LA.; Smith, S. GeoZui3D: data fusion for interpreting oceanographic data; OCEANS 2001, MTS/IEEE Conference and Exhibition; Honolulu, HI. Washington, DC: IEEE Computer Society; 2001. p. 76-84.

12. Arsenault, R.; Ware, C.; Plumlee, M.; Martin, S.; Whitcomb, LL.; Wiley, D.; Gross, T.; Bilgili, A. A system for visualizing time varying oceanographic data. OCEANS 2004; Kobe, Japan: IEEE Computer Society; Washington, DC: 2004. p. 743-747.

13. Andrienko N, Andrienko G, Voss H, Bernardo F, Hipolito JU, Kretschmer U. Testing the usability of interactive maps in common GIS. Cartography and Geographic Information Science 2002;29:325–342.

14. Slocum T, Cliburn D, Feddema J, Miller J. Evaluating the usability of a tool for visualizing the uncertainty of the future global water balance. Cartography and Geographic Information Science 2003;30:299–317.

15. Suchan, TA. Usability studies of geovisualization software in the workplace; National Conference for Digital Government Research; Los Angeles, CA. Marina del ray, CA: Digital Government Research Center; 2002. p. 315-320.

16. Montello, D.; Fabrikant, S.; Ruocco, M.; Middleton, R. Testing the first law of cognitive geography on point-display spatializations. (COSIT '03), Lecture Notes in Computer Science; Proceedings, Conference on Spatial Information Theory; Ittingen, Switzerland. Berlin: Springer Verlag; 2003. p. 316-331.

17. Haklay M, Tobon C. Usability evaluation and PPGIS: towards a user-centered design approach. International Journal of Geographical Information Science 2003;17:577–592.

18. Edsall RM. Design and usability of an enhanced geographic information system for exploration of multivariate health statistics. Professional Geographer 2003;55:605–619.

19. Nielsen, J. Usability Engineering. Academic Press Inc.; Boston, MA: 1993. p. 362

20. Norman, D. The Design of Everyday Things. Vol. 272. Basic Books; New York: 2002.

21. Ericsson, KA.; Simon, HA. Protocol Analysis: Verbal Reports as Data. MIT Press; Cambridge, MA: 1993. p. 443

22. Robinson, AC. Geography. The Pennsylvania State University; University Park, PA: 2005. Assessing geovisualization in epidemiology: a design framework for an exploratory toolkit.

23. Robinson AC, Chen J, Lengerich EJ, Meyer HG, MacEachren AM. Combining usability techniques to design geovisualization tools for epidemiology. Cartography and Geographic Information Science 2005;32:243–257. [PubMed: 19960106]

24. Morgan, DL.; Krueger, RA.; King, JA. The Focus Group Kit. Vol. 692. Sage Publications; Thousand Oaks, CA: 1998.

25. Carr D, White D, MacEachren AM. Conditioned choropleth maps and hypothesis generation. Annals of the Association of American Geographers 2005;95:32–53.

26. Iacopetta B. Are there two sides to colorectal cancer? International Journal of Cancer 2002;101:403–408.

27. Hopenhayn C, Moore DB, Huang B, Redmond J, Tucker TC, Kryscio RJ, Boissoneault GA. Patterns of colorectal cancer incidence risk factors and screening in Kentucky. Southern Medical Journal 2004;97:216–223. [PubMed: 15043326]

28. Howard D, MacEachren AM. Interface design for geographic visualization: tools for representing reliability. Cartography and Geographic Information Systems 1996;23:59–77.

29. Neuendorf, KA. The Content Analysis Guidebook. Sage Publications; Thousand Oaks, CA: 2001. p. 308

30. Anselin L, Syabri I, Kho Y. GeoDa: An introduction to spatial data analysis. Geographical Analysis 2006;38:5–22.

31. Robinson, AC.; Chen, J.; Lengerich, EJ.; Meyer, HG.; MacEachren, AM. Combining usability techniques to design geovisualization tools for epidemiology; Auto-Carto Conference; Las Vegas, NV. Gaithersburg, MD: Cartography and Geographic Information Society; 2005. p. 1-16.

32. Crampton J. Interactivity types in geographic visualization. Cartography and Geographic Information Science 2002;29:85–98.

33. Brodbeck, D.; Girardin, L. Information Visualization Symposium 2003 (Seattle, WA). IEEE; Washington, DC: 2003. Visualization of large-scale customer satisfaction surveys using a parallel coordinate tree; p. 197-202.

34. Roberts, J. Exploratory visualization with multiple linked views. In: Dykes, J.; MacEachren, AM.; Kraak, MJ., editors. Exploring Geovisualization. Vol. 730. Elsevier; Amsterdam: 2005.

35. Guo D. Coordinating computational and visualization approaches for interactive feature selection and multivariate clustering. Information Visualization 2003;2:232–246.

36. Shneiderman, B.; Plaisant, C. Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison-Wesley; Boston, MA: 2005. p. 672

37. Kang, H.; Plaisant, C.; Shneiderman, B. New approaches to help users get started with visual interfaces: multi-layered interfaces and integrated initial guidance; Digital Government Research Conference 2003; Boston, MA. Marina Del Ray, CA: Digital Government Research Center; 2003. p. 1-6.

38. Saraiya, P.; North, C.; Duca, K. IEEE Symposium on Information Visualization 2004 (Austin, TX). IEEE Computer Society; Washington, DC: 2004. An evaluation of microarray visualization tools for biological insight; p. 1-8.

39. Muntz, RR.; Barclay, T.; Dozier, J.; Faloutsos, C.; MacEachren, AM.; Martin, JL.; Pancake, CM.; Satyanarayanan, M. IT Roadmap to a Geospatial Future, report of the Committee on Intersections Between Geospatial Information and Information Technology. Vol. 136. National Academies Press; Washington, DC: 2003.
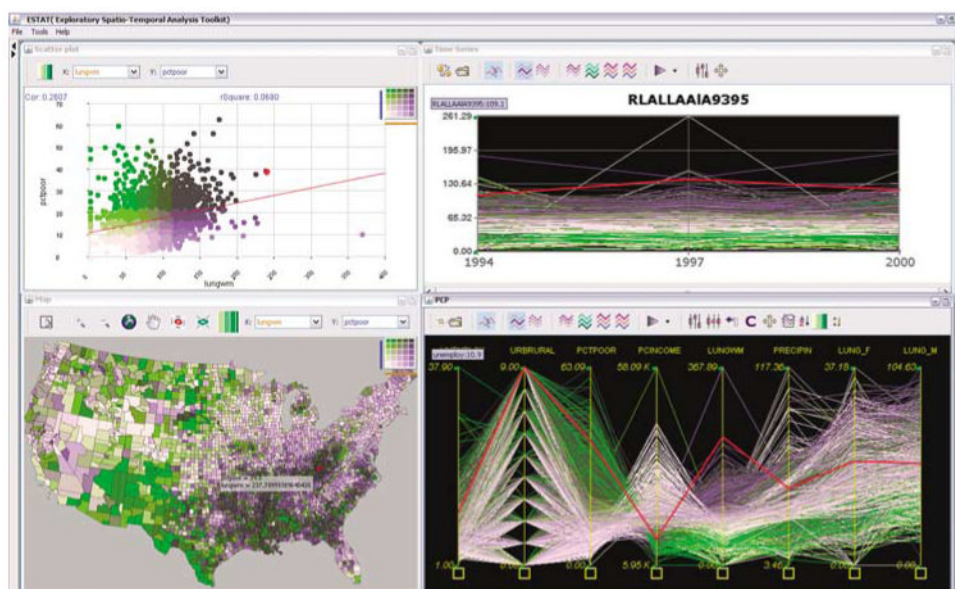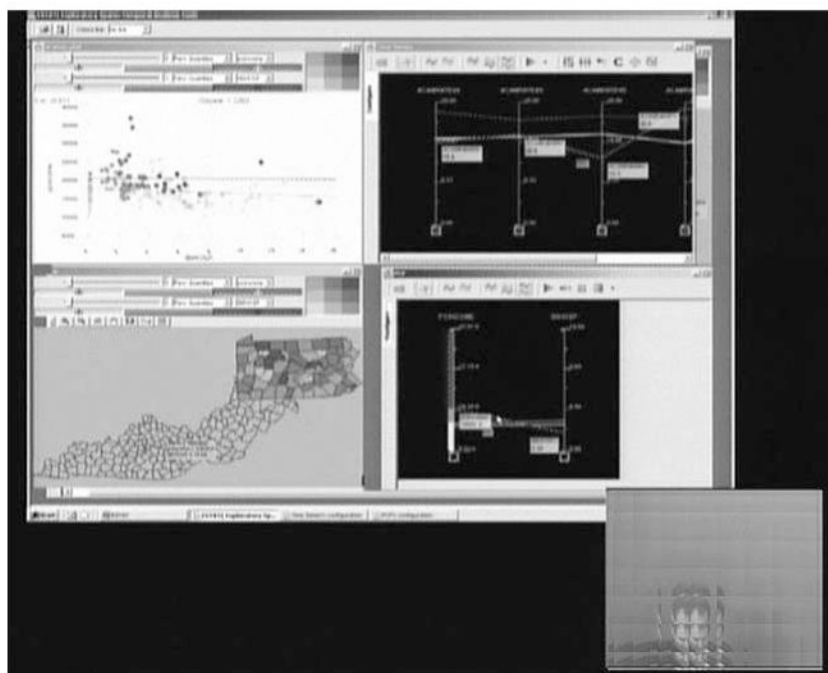
Figure 1.
The ESTAT application features a scatterplot (upper left), bivariate map (lower left), time series graph (upper right), and parallel coordinate plot (lower right). The relationship displayed in the map and scatterplot is a bivariate combination of lung cancer mortality (on the green axis) and the percent population in each county that is living under the poverty level (on the purple axis). The scatterplot shows a weak positive correlation, while the map reveals there are areas of spatial correlation among the high-high counties in Appalachia and the deep south.

Figure 2.
A summary of the formal evaluation activities that have been conducted to date with ESTAT.

Figure 3.
Sample video capture from NCI task analysis sessions. Here the user is examining colon cancer incidence data in Pennsylvania and a number of socioeconomic covariates. The portion of the frame showing the user is distorted to preserve confidentiality.
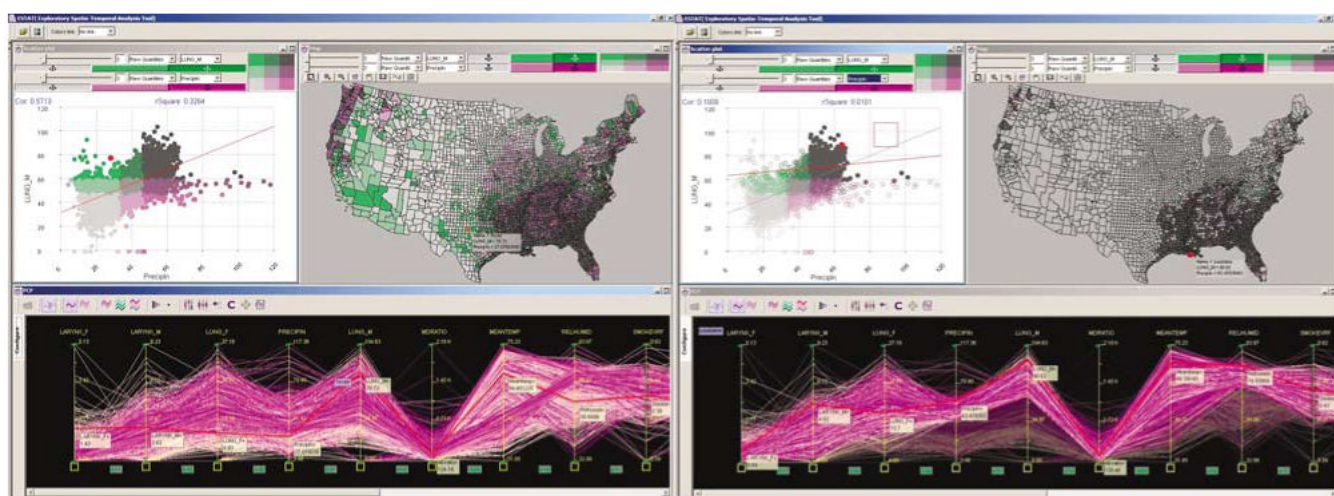
Figure 4.
Task one had participants explore the hypothesis that lung cancer mortality is correlated with mean annual precipitation. The first panel shows the starting point once data has been loaded, and the second panel shows a typical end-result, showing that the top quantile selected in both variables has a strong regional pattern.
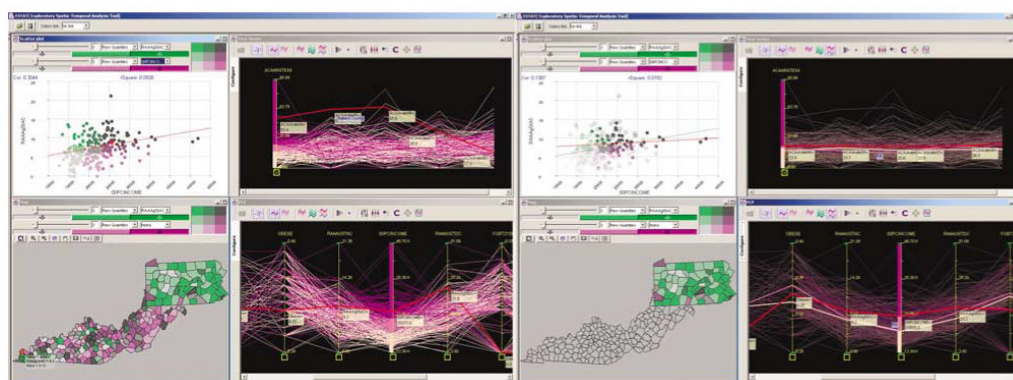
Figure 5.
Task two had participants explore colon cancer incidence data (in this case no initial hypothesis was suggested). The first panel shows ESTAT after data has been loaded. The second panel shows a typical end-result, highlighting Pennsylvania as the state with the highest rates of ascending colon cancer, as well as the most affluent state of the three.

Coding scheme examples:

Tools

"Age adjusted.. oh good so one thing I'm seeing is I can't… I can't scroll this way on this box <gestures over the data description field>…"

"Okay… so these are tertiles? Okay.. <types class # in box> …just to see how it changes."

Application

"Okay… so raw quartiles… quartiles… so… I guess I don't know the difference between the meaning of those. And there's no place to go for help to find the definitions of those?"

"I don't know what I did here… it looks different… it looks different than the graph I got before."

Analysis

So I wouldn't say there's a connection based on the map.. it's quite different from the correlation I mean.. at least for the men it was a little stronger and I'm looking at men here on the map anyway. So that changes my interpretation on the national scale <laughs> if we're just looking at the correlation, but anyway."
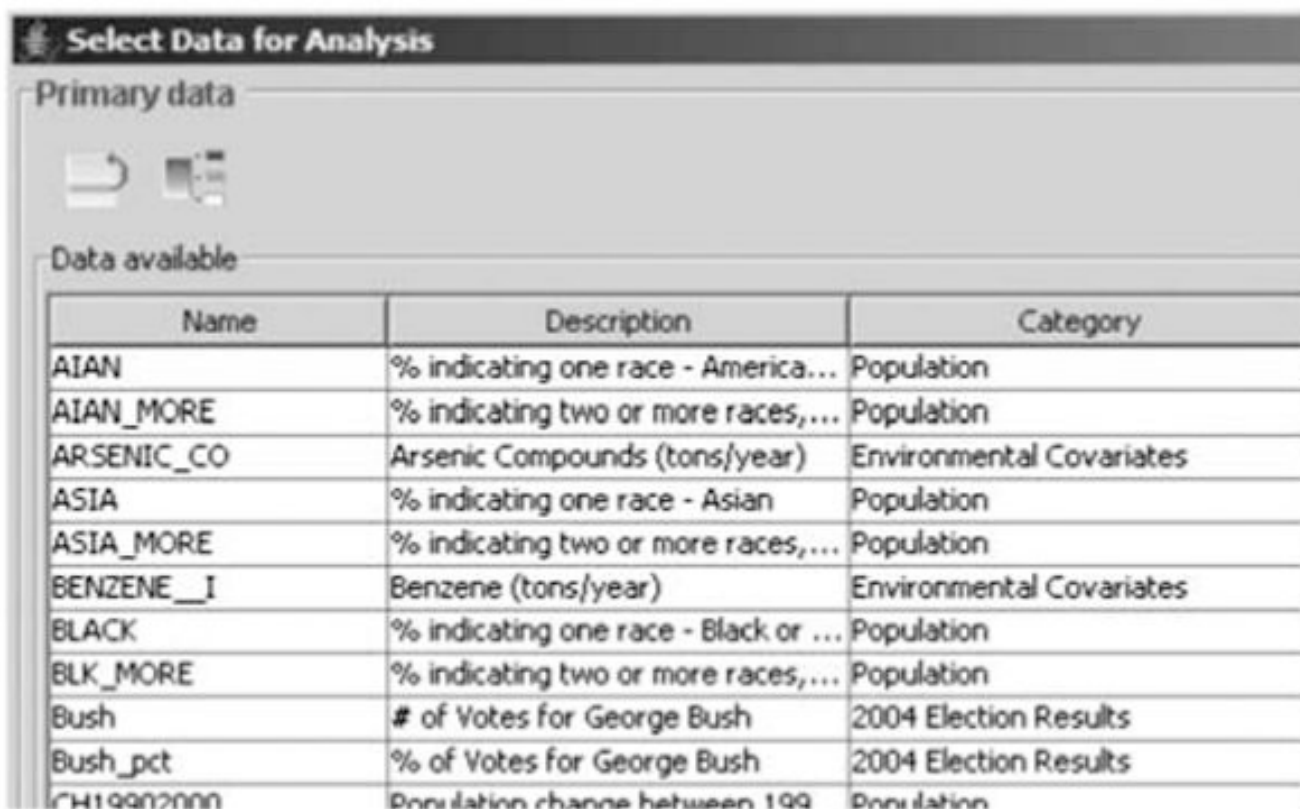
"Okay.. I'll change it to females <changes scatterplot/map variables>.. they have even less correlation.. and the map is.. you can just look at it and see that it's mostly purple or mostly green depending on where it is."

Externalities

"Because I think a lot of the datasets that you might use with this are things that might be part of an ongoing collection.. um.."

"I don't know.. for actual, for making a decision about whether two things are related, I'm more comfortable looking at this correlation or the map. Um."

Figure 6.
Examples of the transcript coding scheme, as applied to representative statements.

Figure 7.
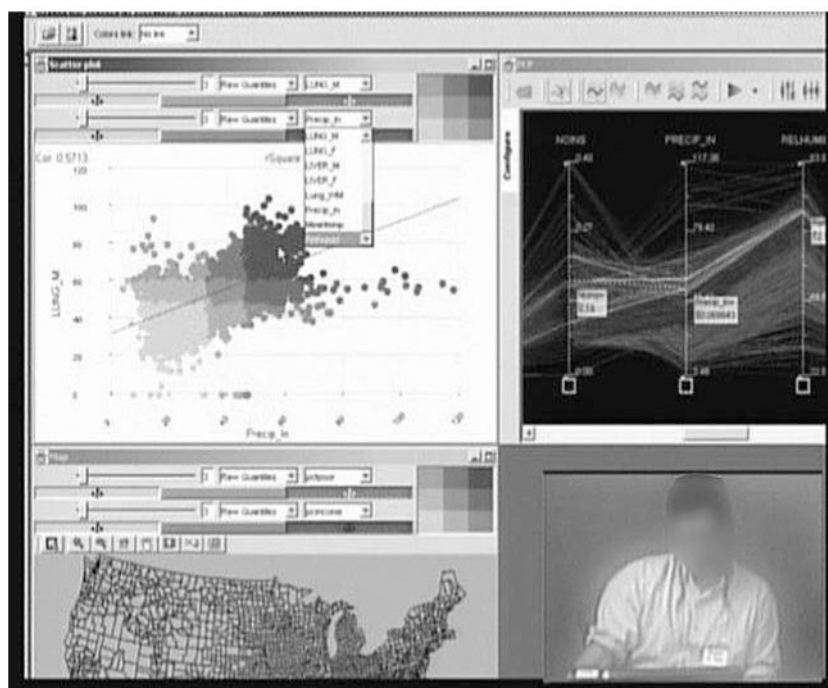The tools shown here allow users to promote all variables in a category to the top of the list (left icon), and to sort all of the variables by categories (right icon).
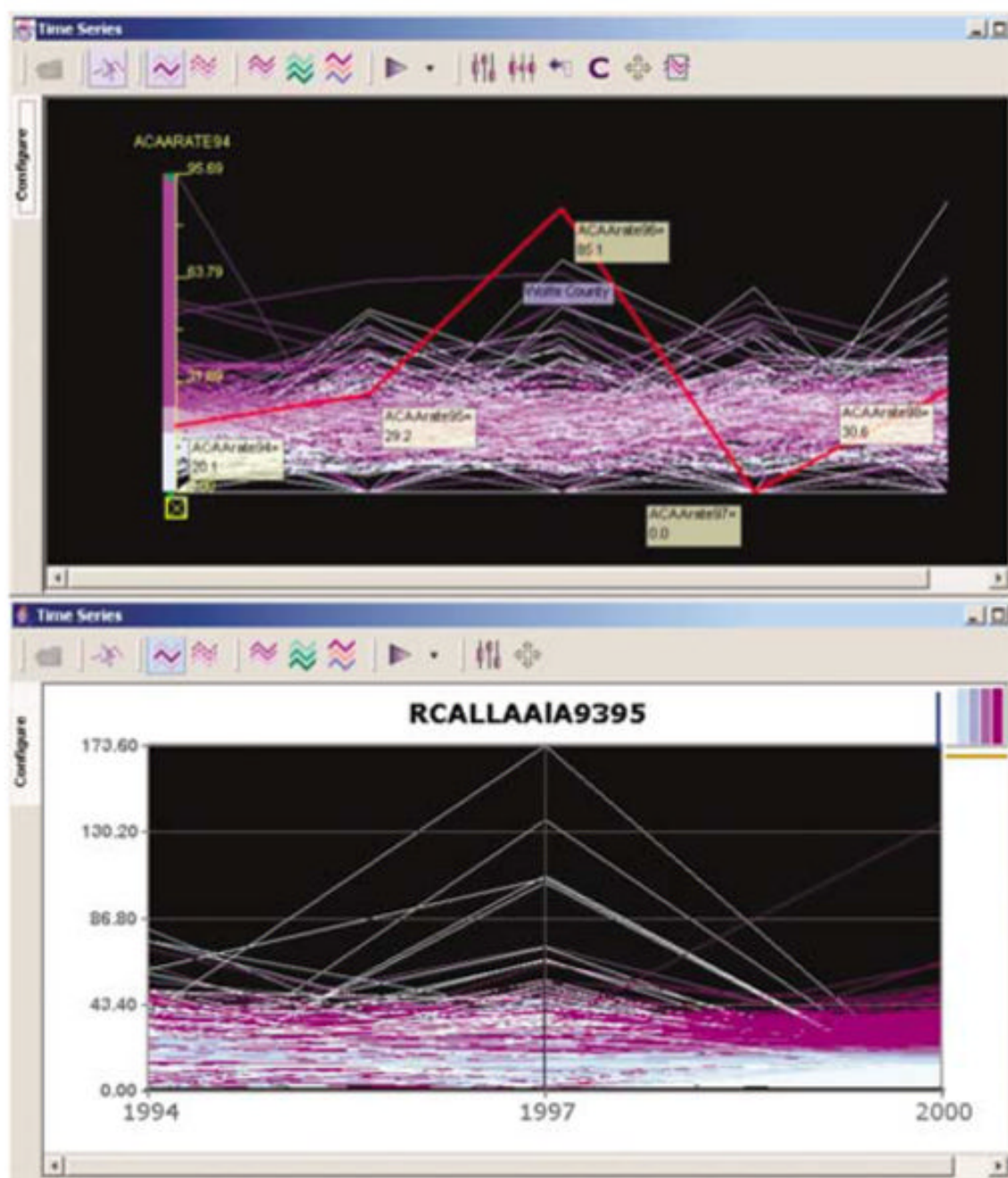
Figure 8.
Capture of P3 using ESTAT as he describes his thoughts about the first task. He interacted exclusively with the scatterplot during the statement shown above. The regression line in the scatterplot indicates a correlation of 0.57 between lung cancer mortality in males and mean annual precipitation.
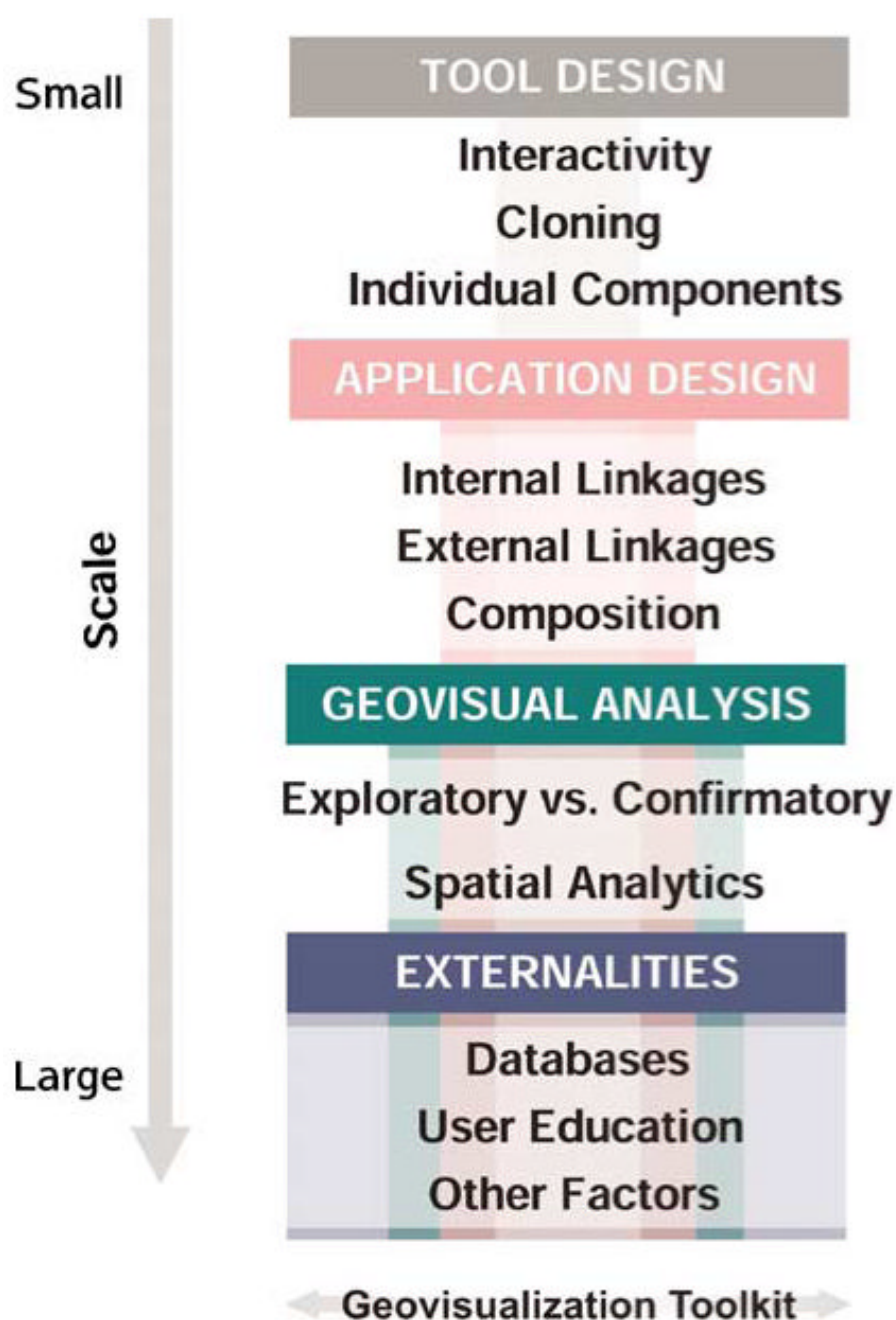
Figure 9.
Design framework graphical summary.

Figure 10.
The old time series graph (top) and the new time series graph (bottom).