Knowledge Precepts for Design and Evaluation of Information Visualizations

Robert A. Amar and John T. Stasko, *Member*, *IEEE*

Abstract—The design and evaluation of most current information visualization systems descend from an emphasis on a user's ability to "unpack" the representations of data of interest and operate on them independently. Too often, successful decision-making and analysis are more a matter of serendipity and user experience than of intentional design and specific support for such tasks; although humans have considerable abilities in analyzing relationships from data, the utility of visualizations remains relatively variable across users, data sets, and domains. In this paper, we discuss the notion of analytic gaps, which represent obstacles faced by visualizations in facilitating higher-level analytic tasks, such as decision-making and learning. We discuss support for bridging these gaps, propose a framework for the design and evaluation of information visualization systems, and demonstrate its use.

Index Terms—Information visualization, visualization techniques and methodologies, theory and methods.

1 Introduction

The modern line of thought on effective presentation of information, espoused strongly by Tufte and others, is that good data speak for themselves [28]. In this sense, Tufte is mainly discussing the creation of static presentations built to convey a message around a particular selected subset of data. Information visualization has grown up around this principle, with the added charge of exploring the benefits of interaction with such displays.

Shneiderman's mantra of "Overview first, zoom and filter, details-on-demand" [22] nicely summarizes the design philosophy of modern information visualization systems, including better-known commercial tools such as Spotfire (2D/3D scatterplots) [27], Eureka (tables with fisheye views and value bars, now the Inxight Table Lens) [14], and InfoZoom (tabular zooming and overview browser) [12]. Beginning with graphical and tabular constructs, these systems provide broad overviews of data sets, support selection and examination of individual data, and provide facilities for dynamic query.

While most recent work on the design and evaluation of information visualization systems typically centers on faithful correspondence of representation to data, there remains uncertainty about current systems' ability to adequately support decision making for three reasons we shall discuss separately: limited affordances, predetermined representations, and the decline of determinism in decision-making.

1.1 Limited Affordances

The operations afforded by many visualization systems are equivalent to very simple database queries. The operations at which these systems excel tend to be those that their default displays and dynamic query interactors afford: simple sorting, filtering, and approximate two-dimensional correlation. A recent study by Kobsa finding that users achieved only 68-75 percent accuracy on simple questions involving some common commercial systems indicates that even these operations have room for improvement [16]. While such operations can be useful for initial exploration of data sets, decision makers are beginning to rely more and more on macro-level, statistical properties of data sets, as we will discuss below.

1.2 Predetermined Representations

The representations employed by common visualizations are not particularly agile, supporting the formation of simplistic, static cognitive models from elementary queries on typically historical, cross-sectional data. If a user's visualization software supports scatterplots, but a contour map is really desired or needed, then a different package must be used. Recently, a number of visualizations that address a specific domain or problem area have emerged ([9], [24], and [29] being examples from the InfoVis '03 Symposium); while they can be very effective, they raise the question of whether each new domain requires a new visualization.

1.3 Decline of Determinism in Decision-Making

Finally, and most importantly, we live in a world that is not only dominated by information, but uncertainty. A growing number of business schools are shying away from information-centric, deterministic management practices; the new managerial "science" is statistical process control [8], with philosophies such as Six Sigma marking an emphasis on managing risk, especially with respect to a growing trend in lowering variability [21].

There is a growing belief that organizations do not resemble mechanical systems so much as holistic organisms, constantly self-organizing and reorganizing to deal with change. According to Freedman:

> In a sense, managers are in a position rather similar to that of pre-chaos natural scientists. They think they understand the

• *The authors are with the College of Computing/GVU Center, Georgia Institute of Technology, Atlanta, GA 30332-0280.*
E-mail: {bob, stasko}@cc.gatech.edu.

Manuscript received 20 Sept. 2004; revised 4 Dec. 2004; accepted 23 Dec. 2004; published online 10 May 2005.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org, and reference IEEECS Log Number TVCGSI-0114-0904.

> relationships between cause and effect in their organizations. But in fact, the links between actions and results are infinitely more complicated than most managers suspect.... As a result, managers are prisoners of the very systems they are supposed to manage. They understand neither the underlying dynamics of these systems nor how to influence those dynamics to achieve organizational goals. [8]

Most information visualization systems do not deal with the notions of uncertainty in data and interlinked causes and effects very well. To be fair, a system can only be as good as the data one provides to it and many systems are optimized for illustrating a few select relationships on a smaller scale. However, data analysts are often interested in complex relationships, especially ones that are not immediately apparent.

In this paper, we examine how these and other limitations in information visualization systems comprise analytic gaps between current systems and higher-level analysis tasks such as learning and decision making. We posit a set of specific knowledge precepts to aid both design and evaluation with respect to higher-level analysis tasks. Finally, we analyze a number of existing systems according to these precepts to identify both where current systems do well with respect to facilitating analytic processes and where there may be opportunities for improvement.

This paper is an expanded version of our InfoVis 2004 conference paper [2]. In that paper, we had conceptualized the design and evaluation framework as a set of "knowledge tasks" to be performed by a visualization designer in support of users performing higher-level analytic tasks such as those listed above. Here, we have clarified our terminology, choosing instead to express the tasks as higher-level knowledge precepts to emphasize the dual role of these precepts in designer and user activities, and expanded our definitions of the analytic gaps that motivate the knowledge precepts.

2 Identifying the Analytic Gaps

2.1 Representational Primacy

primacy (n.)—the state of being first or foremost

The status quo of information visualization is one concerned primarily with what is being visualized, letting designer intuition and user knowledge bridge the gap between the data and its use in higher-level knowledge tasks. As Tufte encourages, "above all else, show the data" [28]. Studies such as Kobsa's [16] test more how well users can unpack the representation of individual data than how users actually discern any higher-level trends or implications of the data sets. This pursuit of faithful data replication and comprehension is what we call *representational primacy*.

What we argue here is that representational primacy can be a limiting notion, perhaps focusing on low-level tasks that do not map well to the true needs and goals of users. Of course, good collection and presentation of data are clear precursors to visualizations of any usefulness. Nor does representational primacy represent insensitivity to users or their needs; rather, it probably represents uncertainty as to how to best support those needs. Technologists have a long history of sharing information and building tools useful to their communities of practice [4]. However, it is not clear that information visualization will be more than a "gee whiz" tool of occasional value to users in general if its use in more analytic thinking is not considered.

2.2 Taxonomies of Tasks: The Status Quo

Desiring to provide a more solid foundation for the design of information presentations, several researchers have developed taxonomies of user tasks that visualizations should facilitate. A significant portion of this taxonomic work concerns itself with developing common user task-related dimensions upon which one can specify requirements and/or desiderata to automated presentation designers, such as SAGE [19], IMPROVISE [32], and BOZ [5].

Roth and Mattis [18] describe a taxonomy of information characteristics used to inform presentation design within the SAGE tool. While much of the taxonomy presented deals with static characteristics of the data, one of its dimensions deals explicitly with user information-seeking goals. Roth and Mattis use two characteristics to deal with such goals: display functions, which vary presentation of a data set based on whether users desire exact value lookup, comparison, or correlation, and distribution functions, which specify how to distribute sets of related information within the presentation.

Wehrend and Lewis create a matrix of representation subproblems that correspond to a particular combination of an object type, such as scalar or vector, and a user cognitive task, such as correlation, distribution, or point identification [30]; the authors identify 11 such user tasks based on a literature search. They then populate the matrix with representation techniques to create a mapping between techniques and problems.

In [32], Zhou and Feiner examine techniques for automatically creating multimedia presentations based on user goals. The authors group high-level presentation goals into two intents: "inform," which deals with elaboration and summarization of data, and "enable," which deals with data exploration and derivation of relationships. They then refine the Wehrend and Lewis operations into visual tasks which are organized by their visual accomplishments (low-level user or presenter goals) and visual implications (what visual capabilities are called upon in the attainment of the visual accomplishments). Each presentation intent maps to visual tasks that achieve it; for instance, the intent "enable-compute-sum" has correlate, locate, and rank.

A common theme in all of these taxonomies is that user goals are thought of as static and explicitly treated only as far as they map into low-level visual tasks. While such tasks are essential, they do not provide a firm basis for supporting the kinds of knowledge-making activities that people seek to perform every day. These taxonomies provide opportunities to improve representations, but do not escape representational primacy.

2.3 The Gaps between Representation and Analysis

A desire to go beyond representationally primal systems has existed for decades, as early as Bertin's assertion in 1977 that "in decision-making the useful information is drawn from the overall relationships of the entire set" [3]. In 2002, Johnston even went so far as to say information visualization was the wrong primary tool where the formation of

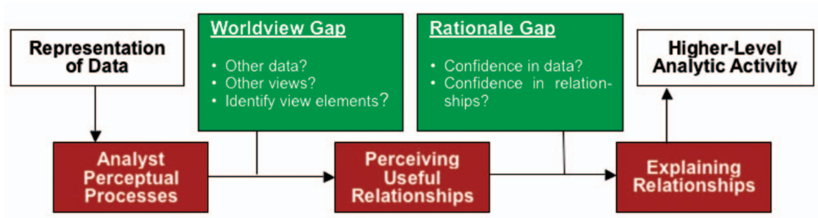

Fig. 1. A model for how bridging the Worldview and Rationale analytic gaps helps facilitate higher-level analytic activity.

explanatory or correlative models was the desired outcome, and asserted a need for "model visualization" rather than "data visualization" [15].

One logical end to this line of thought is to build systems that are "black boxes," in which we input our data and out comes "the answer." However, it is widely viewed as irrational and unethical to trust an important decision to a "black box" system as the rationale for such a decision is obscured and the responsibility for its consequences difficult to allocate. Therefore, we echo the recent arguments of Shneiderman for combining tools such as data mining with information visualization [23] to provide user control.

Shneiderman mainly argues for using data mining to identify time series trends as well as possible correlations for users to explore. We wish to go one step further to what might be called a "white box" approach: systems that promote the generation of higher-level knowledge about a domain that results in justifiable actions. This is certainly a lofty goal which a single system or framework would find difficult to address; however, it is our hope to problematize some of the difficulties visualization systems encounter in such knowledge-making.

We group these issues into two major categories; as these represent distances that must be bridged between current systems and more analytical systems, we call these analytic gaps. Fig. 1 posits a model for analytic activity and frames the two gaps in relation to each other.

2.3.1 The Worldview Gap: Show the Wider Perspective

We define the *Worldview Gap* as the gap between what is being shown and what actually needs to be shown to draw a straightforward representational conclusion for making a decision. Elements of this include: showing appropriate data, using appropriate views to depict the data, and showing relationships clearly.

Although extremely careful data collection and graphic design can indeed create situations where the data indeed speak for themselves, in practice, representation primacy often fails due to imperfect data collection and inexperienced presentation design. Tufte ranks some of the US's most revered journalistic information sources, such as *The New York Times* and *The Wall Street Journal*, as having low graphical sophistication and provides a litany of examples of graphics that decorate numbers rather than actually elucidate relationships between variables [28]. While many information visualization systems are more sophisticated, providing graphical views of correlation and statistical summarization functions, they do not take full advantage of the powerful tools statistics has to offer. While correlation is a gateway to causation, the nature and usefulness of any visualized correlation is uncertain as the true explanatory variable(s) may lie just outside the reach of the data; for example, do family income levels explain standardized test performance, or are the two merely found together?

Nor is it clear that one representation fits all; although scatter plots and graphs facilitate certain comparisons for certain kinds of data, effective representation design remains decided on a case-by-case, domain-by-domain basis. Contrast this with the well-traveled tension of the power of defaults. Kobsa found that Spotfire users tended to use the default scatter-plot visualization in solving problems, even when using a bar chart or histogram representation would have been a better fit [16]. This indicates that representational affordances of a visualization (which, as we have argued, are usually limited) strongly influence what users do with it.

Systems that bridge the Worldview Gap not only indicate the useful relationships among data, but also indicate useful representations and their limits.

2.3.2 The Rationale Gap: No "Black Boxes"

We define the *Rationale Gap* as the gap between perceiving a relationship and actually being able to explain confidence in that relationship and the usefulness of that relationship. Elements of this include: confidence in data, highlighting uncertainty in data or data aggregation, and understanding the effects of change.

Systems built under representational primacy assist in the perception of relationships, but very often fail to elucidate the strengths of these relationships and the confidence in these relationships. As a simple example, comparing averages in a visualization tool is misleading unless you know something about the populations from which the averages came and, thus, your confidence in the actual difference in averages. As a tool incorporates a wider range of techniques, this problem compounds itself.

Indeed, typical implementations of business intelligence software have proven to be overly complex and require too much specialist intervention; the end result is not analytic clarity but an endless stream of reports [7]. Systems that bridge the Rationale Gap not only provide accurate, clear answers, but instill in users identifiable rationale about the kinds of decisions that can be made through their use.

3 Example Analytic Gap Scenarios

In order to provide further grounding for these gaps and how existing systems can fall into them, we provide two example scenarios.

3.1 Example: Sports Analysis

Consider being the general manager of a sports team, with the responsibility of trading and acquiring personnel to produce the best results. (In fact, many people live this dream daily through fantasy sports competitions.) Analyzing a data set of sports statistics for some given year for leaders in statistical categories is fairly straightforwardly done using current visualization tools and dynamic query operations such as sorting. With a tool that provides aggregation functions, one can even relatively quickly compare the actual performance and payrolls of whole teams across the year, such as exists in individual leagues or in the sport as a whole.

All of this is useful in making some intuitive sense out of the data given; it can be especially useful in spotting anomalies such as extremely poor or good performers or extremely high team payrolls. Still, there are two major problems.

First, while most tools visualize correlations and simple relationships, they fail to provide indications as to which relationships or combinations of relationships most strongly suggest the attainment of a certain performance metric, such as win percentage or offensive effectiveness, falling into the Worldview Gap and leaving users to use their own intuition as to what aspects of the data set are most useful. Confounds in correlation of variables are especially troubling when decisions involve a lot of money, such as those about sports personnel movement. Second, any intuition we may develop about the data set is hard to transfer away from the tool; we may be able to see correlations for two or three variables at one time, but what we really desire is a plug-and-play "causal model," especially for predictive actions such as determining future performance of certain players. Unfortunately, information visualization systems provide little to no support for the formulation of predictive models, let alone a clear explanation as to how such a model might be constructed, running headlong into the Rationale Gap.

Possibly even more troubling is that we cannot really use a visualization tool to apply any real-world constraints, such as economic constraints; while we can dream about the sorts of teams we can put together and even get a superficial sense for how much such teams will cost, we cannot easily reason about how to achieve such an outcome in reality, such as managing money to be committed to players in the future and coping with the effects on the existing organization and personnel.

Such forward-looking prediction is arguably out of the domain of a representational visualization tool. Still, we believe it is not implausible for at least some of the analytic processes involved to be translated into the perceptual domain, offering a viable and accessible complement to data mining tools and spreadsheets.

3.2 Example: Managerial Decision-Making

In his book *The Fifth Discipline*, Senge describes a case study of a fictional company called WonderTech, which began by growing quickly but eventually collapsed under huge cycles of alternating high and low demand. The end result was due to a vicious circle of needing to improve sales but not having the capacity to keep up with sales when they did improve; as a result, the fixed investments in manufacturing increased, but sales failed to stay consistently high enough to support an increasingly expensive infrastructure. [20].

Here is an instance when, presumably, the managers of WonderTech had a multitude of numbers available to them and possibly even saw the cyclic nature of their sales figures and a growing fixed cost of manufacturing, but either failed to see the basic feedback process, failed to see a way out of the feedback process, or were too occupied with short-term solutions to get an accepted long-term solution in place, such as a commitment to rapid delivery [8]. Most visualization tools would support a time-series view of sales and financials which would go far in elucidating that there was a problem. However, it would take a miracle in the data set to show growing order fulfillment times and an as-of-yet nonexistent capability to show that reducing these fulfillment times could result in a better long-term ability to support sales, an example of the Worldview Gap.

3.3 Prototypical Analysis Tasks

These example scenarios show how current information visualization systems may fail to provide the support needed to an analyst. The tasks being performed by the analyst in each scenario are higher-level than the "correlate," "identify," and "distribute" tasks identified by the taxonomies discussed previously ([18], [30], [32]). Under representational primacy, a system designer provides and refines support for these and other low-level visual tasks. A reasonable analogous approach to designing systems to bridge analytic gaps is to identify high-level user analysis tasks that are not evident in the existing taxonomies, but that are typical of what real people want to accomplish using visualization tools. Below, let us consider a few such higher-level tasks. Our goal is not to present a taxonomy of these tasks (especially since they may contain analytic subprocesses), but, rather, to illustrate some common high-level processes and how they fall into the analytic gaps.

Complex decision-making, especially under uncertainty. The higher-level analysis tasks performed by people every day are often done in service of concrete decision-making or planning processes, such as budgeting, strategizing, and cost-benefit analysis. In the spirit of the sports team management example from earlier, consider the seemingly simple task of deciding whether to trade players with another team. It is far from straightforward to understand the expected collective performance of subsets of players and the costs and benefits to teams of making personnel changes.

These kinds of analyses require flexibility in elucidating the variables involved in the relationships, which requires bridging the Worldview Gap. Techniques such as sensitivity analysis answer questions relating to the Rationale Gap.

Learning a domain. A person unfamiliar with Formula One car racing might examine a rich data set containing all race results in history. The person could learn about the different drivers that have been most successful, where the races are held, the types of cars used in these races, whether there have been any fatalities in the races, and so on. Such a data set may teach a great deal about the topic of Formula One racing and examination may suggest data that could be added to make the data set even more informative.

In general, exploration of particular data sets can reveal a lot about the general discipline or phenomena that the data sets describe. It can also suggest elements outside the data set that further elucidate the domain, a clear appeal to the Worldview Gap.

Identifying the nature of trends. Casual examination of data or certain data attributes, given the right representation, can reveal trends. Trends are sometimes as simple as correlation or time-series expansion, but, for more complicated, aperiodic, or freeform trends, straightforward visual representations can mislead or obscure.

Many financial and scientific professionals are typically concerned with trending, especially in the short term. One particularly vivid instance is that of the "technical" stock investor, who makes buying and selling decisions based entirely upon the presence of certain signal patterns in the price of a given stock and the market as a whole. The Worldview Gap asks an analyst to consider where patterns might be found and the Rationale Gap asks how much confidence one might have in such patterns (e.g., stability, difference from background variation).

Predicting the future. Forecasting what will occur in the future is a common activity in many different walks of life and relies heavily on an understanding of the past. To make a forecast, an analyst must gather as much current data as possible and then examine the data in the context of past data and outcomes.

Consider how weather forecasters examine data about current meteorological conditions including temperatures, winds, pressure systems, etc., in order to predict future weather. Investment analysts want to buy stocks that will have the greatest appreciation in the future. These analysts typically gather much data about the fundamental health of companies, including financial situations, current operating practices, and revenue streams. A Worldview Gap perspective considers which variables are most important and how to find those variables and a Rational Gap perspective considers how strongly we can believe in the predictions themselves.

As we have described, these analysis tasks fall into the analytic gaps in various ways. The Worldview Gap encompasses shortcomings in domain knowledge as well as data sufficiency and the Rationale Gap demands the clear identification of decision criteria and confidence. In order to fulfill the goal of actionable knowledge, both gaps must be filled.

4 Bridging the Analytic Gaps: Knowledge Precepts

Many higher-level tasks and task variations exist. We use the idea of the analytic gaps to identify common components of analytic processes in which visualization can provide direct benefits. Here, we wish to advance our framework of knowledge precepts to guide both users and designers to solutions which help bridge the analytic gaps. These are just a few precepts that we have recognized; others most certainly exist. We classify each precept according to which analytic gap primarily motivates it, although overlap is possible.

4.1 Worldview-Based Precepts

Many precepts we will describe here indirectly support the formulation of a strategy for browsing a visualization when they provide insights as to what data should be explored to clarify certain relationships or test certain hypotheses.

4.1.1 Worldview Precept 1: Determine Domain Parameters

The attributes of data in a visualization and, thus, the parameters by which data is organized in a visualization communicate both standards of measure within a data set and key parameters for understanding a domain. This communication is not just a dialogue between an analyst and a canonical description of sense-making within a domain, but also between data collectors and analysts.

The very fact that a collection of American baseball scores includes data such as home runs, runs batted in, and slugging percentage indicates that these are parameters considered important (at least by the data collector) and suggests domain-specific measures that require clarification. As well, the relative positive or negative connotations of parameters are not always clear; in American baseball, a batter with a high number of career strikeouts may not be considered a good batter nor a pitcher with a high number of walks and hits allowed. These interpretations, however, are not always inherent in the visualization.

We consider this a Worldview Gap precept because it points the way to formation of more expressive representations. To summarize, *a system can help bridge the Worldview Gap by providing facilities for creating, acquiring, and transferring knowledge or metadata about important domain parameters within a data set*.

4.1.2 Worldview Precept 2: Expose Multivariate Explanation

Most visualization systems support determination of correlation between two or three variables in the limit of representational ability. However, some relationships involve more than three explanatory variables and/or simple transformation of single explanatory variables using logarithms or polynomial relationships [1]. Such correlations, often found in domains such as queuing theory, are not widely handled by typical visualization tools. Also, when correlations expected by theory do not exist, correct interpretation and action usually involves user guidance. In general, while statistics offers methods such as stepwise regression to help automatically determine good explanatory models, mindlessly employing such tools generally yields bad results [1]. Combining these methods with user guidance could result in a very useful facility for data analysts.

In 1990, La Quinta Motor Inns retained the services of academic statisticians who derived a successful mathematical model for the selection of sites for La Quinta inns [1]. The model directly related site profitability to the room rate and inversely related profitability to the population of the state of the site, which both seem reasonable. However, the analysts also found a strong direct relationship between profitability and the number of college students within four miles (possibly surprising) and an inverse relationship between profitability and the square root of the median income of the area. The model explained 51 percent of the variation in profitability, which is respectable in practice;

however, this possibility does need to be raised to a user of the model, who may experience deviations from the results.

The emphasis of this precept is on the discovery of useful explanatory variables, either automatically or manually, which directly bridges the Worldview Gap. To summarize, *a system can help bridge the Worldview Gap by providing support for discovery (whether automated or manual) of useful correlative models and constraints*.

4.1.3 Worldview Precept 3: Facilitate Hypothesis Testing

Users need to test the accuracy of their deductions about a data set. Tools must help users define hypotheses, simulate possible outcomes, and verify the truth of such hypotheses. While we might include statistical hypothesis tests, such as confirmation of expectation (e.g., statistical distribution of results, expected limits of data values) and comparison of averages with certain confidence intervals, this task includes higher-level hypotheses. If a particular region or outcome of interest is found, then hypothesis tests can also become a question of how far and how easily users can operate on that outcome. This analytic process is clearly difficult to support generally across interfaces and representations, but may be useful for specific design decisions.

It should be noted that here we are considering the interactive and iterative nature of users' browsing of a visualization. In fairness, these activities can include deducing causation and other relationships pertinent to the Rationale Gap. We include the activity of hypothesis testing here with the Worldview precepts since it can encompass other activities, based on the expressiveness and completeness of cognitive or mathematical models derived from use of a visualization. To summarize, *a system can help bridge the Worldview Gap by providing support for the interactive formulation and verification of user hypotheses*.

4.2 Rationale-Based Precepts

Users need to be able to relate data sets to the realms in which decisions are being made. For example, analysis of a computational chemistry data set may produce an encoding for a promising lead compound for the design of a drug [6]. Proper visualization of the data set communicates how to modify existing compounds to obtain the promising lead. Also, given a set of criteria, users need to be able to use salient features of data sets to create a description of the realm in general to validate decisions.

4.2.1 Rationale Precept 1: Expose Uncertainty

Some uncertainty is involved in any data set. Is the data set large enough to mitigate any associated sampling error? Are there figures in a data set involving uncertainties, such as population estimates, with associated standard errors or statistically distributed phenomena? An understanding of where values are uncertain and how that uncertainty affects the degree to which a data set can be a source for reliable conclusions is key in statistical process control.

For example, when considering several vendors for a part whose width must be exactly within a specified range, it is important to understand not just the width of the average part produced, but the standard deviation as well (to understand the proportion of unusable parts). Also,

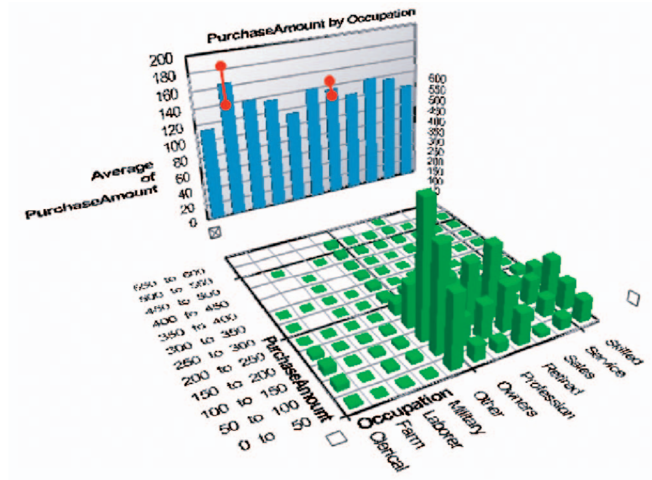

Fig. 2. Error bars (which we have added in red) would be a simple way to increase confidence in the degree of difference between two aggregations. (Picture taken from the SeeIt system by Visible Decisions, Inc.)

when comparing estimated financial figures or poll results, such as in Fig. 2, having a measure of estimation error is crucial to having confidence in the statistical significance of differences observed, especially with small sample sizes.

We consider this a Rationale Gap precept as it relates directly to the confidence one can draw based on correlation or aggregation analysis done within a visualization tool. To summarize, *a system can help bridge the Rationale Gap by exposing uncertainty in data measures and aggregations and showing the possible effect of this uncertainty on outcomes*.

4.2.2 Rationale Precept 2: Concretize Relationships

In the case of correlation, especially when viewed on a scatterplot, perceived relationships are usually easy to describe and quantify. Other representations may suggest relationships or decisions without a clear concretization of the nature of the relationships. This can be particularly problematic in expected value analysis. When the expected payoff or outcome of a decision is a weighted average of the elements of a clearly identifiable discrete distribution (called a risk profile in business), then the outcomes are not so clear and are often surprising when people think in terms of expected values.

This is a Rationale Gap precept in the spirit of being able to rationalize decisions and outcomes based on a cognitive alignment of a perceived relationship with its representational elements. It shares the spirit of Worldview Precept 2 (expose multivariate explanation), but concentrates on end-user presentation rather than discovery. To summarize, *a system can help bridge the Rationale Gap by clearly presenting what comprises the representation of a relationship and presenting concrete outcomes where appropriate*.

4.2.3 Rationale Precept 3: Expose Cause and Effect

When investigating data, there is usually some causation data embedded directly in the data set, as well as effect data that can become clear through iterations of a simulation. Both the isolation of demonstrated causes as well as the discovery of possible effects are important in cognitive model formation. All of this must be done with an understanding of what assumptions have gone into creating the data and thus affect the outcomes inferred. As an example, consider the story of WonderTech we recounted earlier. Some causation can be inferred from time series data of sales and manufacturing costs; a further step would be to be able to investigate the effects of changing certain variables on the outcomes depicted by the data set, such as sensitivity analyses (e.g., the value of an investment as depends on factors such as interest rates or growth predictions).

This precept has elements of the Worldview Gap in it since causation can reach beyond the data set; in categorizing it as a Rationale precept, we consider how causations form the basis of relationships inferred from visualizing a data set and the factors for distinguishing between true causation and mere covariance. To summarize, *a system can help bridge the Rationale Gap by clarifying possible sources of causation*.

5 Employing the Knowledge Precepts

The analytic gaps and knowledge precepts we have proposed so far also form the basis for a design and evaluation framework. In essence, all one need do is apply the knowledge precepts (plus any other higher-level knowledge tasks or precepts one wishes to employ) to a given situation.

5.1 Using the Precepts for Design

When designing a visualization for a new domain or scenario, one can use the set of knowledge precepts in order to systematically:

1. Generate new subtasks for a visualization to support or perform.
2. Identify possible shortcomings in representation or data.
3. Discover possible relationships to highlight or use as the basis for a visualization.

The general idea is to apply each knowledge precept in turn as a user would to each scenario. For example, "Where might I be interested in multivariate relationships?" or "Exactly what is uncertain about this data and how will it affect the outcomes I show?" or even "How will I show the concrete outcomes from this process?"

5.2 Using the Precepts for Evaluation

One can also use these precepts as a form of heuristic evaluation [17] of the pragmatic value of a given visualization simply by evaluating how well the visualization supports the knowledge precepts. The Rationale Gap precepts provide particularly rich opportunities to ask questions both about how actual relationships and outcomes are shown to a user (e.g., must the user infer an outcome from the context of a representation or can a user perform a direct action to see an outcome, such as in a brushing histogram), as well as how confident the user should be in these outcomes relative to any uncertainty inherent in the data set being visualized.

6 Design Example: The Infovis 2004 Contest

While we have mainly discussed quantitative scenarios, such as those found in the financial and scientific domains, the six knowledge precepts here provide a very fruitful way of thinking about visualizations for a decidedly less quantitative scenario: the InfoVis 2004 Contest [11]. The contest, which is to provide visualizations to support questions about the evolution of information visualization as a research area, is based on a data set containing metadata (titles, abstracts, keywords, dates, and references) about articles from the InfoVis conference from 1995 to 2002. Although it is hoped that applying the knowledge precepts sheds new light on possible solutions to contest tasks, we wish to show more that the knowledge precepts provide a systematic basis for thinking about and identifying issues in the data set.

6.1 Worldview Precept 1: Determine Domain Parameters

Clearly, the attributes of the metadata dominate our thinking about the data set. We have already discussed the notion of considering other factors that may come to bear on the data set that might not currently be reflected. Another possibility is to consider how deeply the metadata allow us to make conclusions. Are abstracts enough to relate articles or do we need more text to do the appropriate comparisons? Are references enough or do we need more metadata on what kinds of papers (conference full papers, extended abstracts, technical notes, journal papers, etc.) are citing other papers and being cited?

6.2 Worldview Precept 2: Expose Multivariate Explanation

Let us consider, as a hypothetical approach, the use of a concept map, such as a themescape [31], in which we show two or more researchers' work as regions on that themescape, highlighting areas of overlap with brighter colors to indicate the degree of overlap of the researchers involved. The outcomes highlighted for the user are a two-dimensional projection of a potentially multivariate trend. The important questions for design of a relevant visualization include generating possible multivariate explanations as well as how to communicate the variables' contribution to the overall analysis. For example, one may determine that the trajectory of a researcher on a themescape is determined by a particular correlation with the subject matter of other researchers, dates of publication, and keywords (possibly both author-provided and contest entrant-generated).

6.3 Worldview Precept 3: Facilitate Hypothesis Testing

Even though the contest tasks are mainly qualitative, users may wish to experiment with different classifications or evolution along different dimensions: For example, using research money allocated to areas or a number of people working in an area to show evolution rather than a size-agnostic time-based evolution. Considering the themescape example again, if overlaps are identified in fringe areas, a user may wish to see if that fringe area eventually panned out into anything larger. One may even wish to ask higher-level questions, such as whether or not the development of

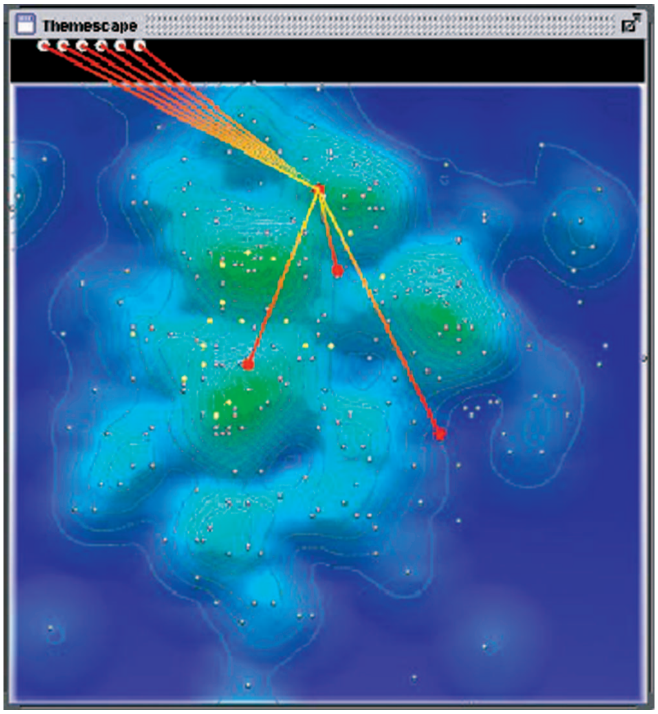

Fig. 3. This themescape variation allows documents with missing metadata, shown as dots in the upper black region, to participate in analysis, such as the reference relationship shown. (Picture courtesy of Nicholas Diakopoulos.)

a particular research area was hindered by or depended upon the development of a different area. Ultimately, for the purposes of the contest, this form of experimentation may be limited, but considering the types and degree of utility of such experimentation may help decide the feature set available to a user.

6.4 Rationale Precept 1: Expose Uncertainty

For this data set, dominated primarily by nominal data, at first glance it seems there is no uncertainty to speak of. However, uncertainty can appear in more forms than standard deviations and measurement errors. If one examines the metadata for completeness, one notices a number of possible sources of uncertainty. For example, author names are sometimes spelled or formatted differently. Paper dates are sometimes exact and sometimes involve a large range of dates. References may be missing or their formats may differ, requiring significant effort for tagging or cleaning.

In other words, being sure of who is who, when is when, and sometimes even what is what is difficult. If any uncertainties cannot be resolved in the process of data cleaning, they must be shown to the user. For example, if it is unclear whether or not "J. Smith" and "J.T. Smith" are the same person, this is an uncertainty, especially given the higher-level tasks contest entrants are asked to support. Fig. 3 shows a possible augmentation to a themescape that shows these uncertainties while still allowing for incomplete data to participate in analytic activity.

6.5 Rationale Precept 2: Concretize Relationships

If we are asked to relate two researchers' work in the field of information visualization, how will we do it? Ideally, a visualization should provide perceptual triggers [26] to show these outcomes. Returning to the themescape example, what is keyed to the brightness of areas of overlap? Is it frequent coauthorship, common mutual referencing, unity in research subject matter, or something else entirely? If there is significant overlap in fringe areas, then does that represent the formation of new research areas or just a coincidence? All of these items could be indicated to the user.

6.6 Rationale Precept 3: Expose Cause and Effect

Here, we can think about possible causes and effects in the field to generate interesting ideas for relationships to highlight. Did one paper spawn off a generation of related papers? Can we identify opposing schools of thought on a topic and their point evolutions in time? Do user studies (tagged externally by other participants) promote new and interesting ideas in the field? Most importantly, what data must we employ to validate this cause and effect? How can a user feel he/she is exploring the data set and knows where the relationships come from, rather than interacting with a "black box"?

7 Evaluation Example: Current Commercial Tools

We can also use the knowledge precepts to reflect upon how commercial tools might or might not be meeting the challenges posed by the analytic gap. Here, we consider the same trio of tools considered by Kobsa in his evaluation [16]: Spotfire, Eureka, and InfoZoom.

7.1 Worldview Precept 1: Determine Domain Parameters

Since these systems are largely data-driven, the tools communicate the domain parameters that are in the data set. Most of the issues here revolve around presentation; for example, Spotfire relegates some data to a smaller window for details-on-demand and Eureka occasionally has problems displaying large labels. The ability to attach annotations or other metadata to domain parameters and present such metadata to the viewer would be advantageous.

7.2 Worldview Precept 2: Expose Multivariate Explanation

Spotfire offers explicit three-dimensional correlation; while Eureka and InfoZoom do not offer explicit correlation, they do use filtering, brushing, and sorting on many different attributes at once. For flexibility and ease of analysis, these systems could provide more tools for correlation, such as nonlinear correlation and correlation to logarithms or polynomial functions of data.

7.3 Worldview Precept 3: Facilitate Hypothesis Testing

In these systems, when items of interest are isolated, their context in the data set as a whole is visible. As mentioned before, InfoZoom does provide powerful derived attributes; in fact, all the tools provide some way of creating at least

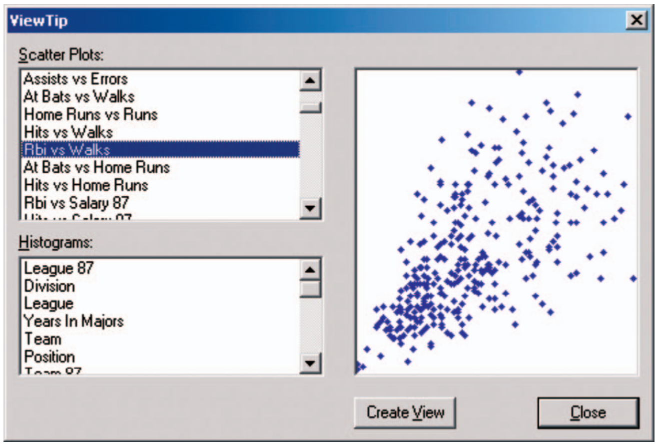

Fig. 4. The View Tips in SpotFire Pro 4.0 allow users to quickly examine possible sources of correlation for further examination.

simple derived attributes, usually based on aggregation functions. However, the tools are not as well suited to time series analysis, which is a common basis for higher-level data analysis and hypothesis tests.

7.4 Rationale Precept 1: Expose Uncertainty

Again, most statistical facilities in these information visualization systems are limited to aggregation and correlation. Spotfire can bin data according to standard deviation and can indirectly show some variations around points, but the explicit treatment of uncertainty is otherwise limited. Eureka and InfoZoom generally display the data as given. None of the programs allow easy comparison of averages within a certain confidence, although InfoZoom's "derived attributes" functionality is programmatically expressive for those who can write programs. Granted, the data provided do not always show uncertainty well; still, uncertainty is not generally part of the data import facilities of these programs and, even if explicit measures of uncertainty were integrated into the data, the data importing facilities would require them to be treated as data rather than metadata.

7.5 Rationale Precept 2: Concretize Relationships

All of these commercial systems can show details-on-demand for a particular item or set of items. As well, when filtering relationships are applied, single items or sets of items can be easily shown and isolated for individual examination. However, close but inexact matches, as well as relationships based on probabilistic links, are harder, if not impossible, to show and isolate. An approach such as the Attribute Explorer [25] can help increase the flexibility of such queries.

7.6 Rationale Precept 3: Expose Cause and Effect

Spotfire provides the "View Tips" functionality, shown in Fig. 4, that highlights interesting correlations for users. Otherwise, users of these systems are left on their own to explore possible correlations. As well, no facilities for sensitivity analysis are provided.

8 Facilitating Analytic Tasks: IN-SPIRE

As a concluding example, we examine a visualization system that addresses some of the analytic gaps set forth in this paper. The IN-SPIRE system ([10], [13], [31]) is a visual tool for processing and mining large amounts of text documents. The major visualization components include automatically generated landscapes similar to themescapes [31] and galaxy-based views for identifying individual documents of interest. IN-SPIRE also includes several interactivie facilities for querying and tracking documents. Besides concordance and text query functions, the system contains a hypothesis manager for keeping track of documents which support and contradict hypotheses.

8.1 Supporting Worldview Precepts

Concordance facilities and the ability to query text similarity in a variety of ways allows the ability to isolate documents that may be thematically related but discuss different subject matter. This contributes to understanding domain boundaries and possible cross-domain relationships, in support of Worldview Precept 1. Although a lot of the burden is on the user to manually explore related documents, IN-SPIRE's classification schemes allow quick visual pruning of the search space.

The ability to control the strength of classification and grouping of documents in the landscape and galaxy representations promotes the possibility of understanding multifaceted relationships and chains of relationships (Worldview Precept 2). If the terms providing the strongest linkages within theme groups are relegated to "outlier term" status, then redistribution on the secondary themes of the documents can provide grounds for further exploration. Again, while the system does not perform the exploration for the user, the major results are organized in a way that allows users to quickly identify groups of interest.

The Hypothesis Manager provides a way of formulating, tracking, and analyzing simple hypotheses, supporting Worldview Precept 3. Users can keep track of which documents support and refute user-specified hypotheses, with special emphasis placed on "diagnostic" documents, which refute one hypothesis and support a different one. IN-SPIRE's time-slicing capabilities can also provide some indirect support for temporal hypotheses, such as those involving the development of themes in the text corpus.

8.2 Supporting Rationale Precepts

The way in which IN-SPIRE categorizes documents helps to make relationships within and between groups of documents concrete, as given by Rationale Precept 2. IN-SPIRE provides clear labels that govern theme peaks and clusters. The ability to specifically exclude words as "outlier terms" provides easy ways within the tool of exploring the different factors making up the document groupings. The density of theme labels is easily adjustable, making identification of subthemes or fringe areas easier.

Time-slicing capability combined with rationalized groupings promotes the determination of origins of selected subject matter lines, helping to promote the discovery of temporal cause and effect, which is Rationale Precept 3. Users can specify groupings (such as papers regarding a

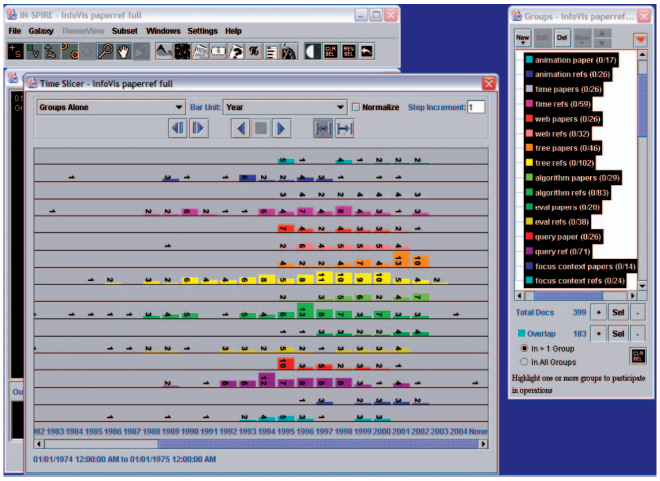

Fig. 5. IN-SPIRE uses horizontal scrolling to navigate time slices of user-defined content groups. (Picture produced at and provided with permission of Pacific Northwest National Laboratory, which is managed and operated by the Battelle Memorial Institute on behalf of the US Department of Energy.)

specific topic or written by a certain author) and easily see their thematic distribution for a cross-section of time in the bar graph or galaxy views. A user can scroll forward and backward through time to see the emergence and formation of these groups. This capability is shown in Fig. 5.

9 Conclusion

In this paper, we have identified the focus of current information visualization systems on representational primacy or the overriding pursuit of faithful data replication and comprehension. We have argued that to become even more useful, a parallel focus on what might be called *analytic primacy* must emerge. Limitations in current systems were classified into one of two analytic gaps: the Rationale Gap, representing the gap between perceiving a relationship and expressing confidence in the correctness and utility of that relationship, and the Worldview Gap, representing the gap between what is shown to a user and what actually needs to be shown to draw a representational conclusion for making a decision. For each gap, we proposed three precepts that suggest ways to narrow or diminish these gaps and then demonstrated how these precepts might be used for systematic design and heuristic evaluation with respect to analytic activity.

We have only begun to identify and catalog the issues surrounding analytic primacy. While we have discussed a number of statistical and analytic phenomena involving correlation and model formation, the precepts apply to related phenomena such as clusters, outliers, and gaps as well (although humans tend to be good at perceiving such patterns and anomalies). As well, the social aspects of analytic primacy potentially reach far beyond our discussion of finding appropriate domain parameters and metadata, extending into widespread collaboration in service of sense-making and decision-making.

While we have primarily concentrated on information visualization, similar challenges and problems exist in other visualization realms such as scientific visualization. In providing a set of knowledge tasks and a framework of knowledge precepts, our intention is to check the status quo of visualization tools with the decision-making processes of the real world. In short, we are asking what more these systems could do to be more useful for decision makers. If, as Tufte asserts, we lack graphical sophistication as a population, then perhaps we need all the help we can get to make sense of the rapidly burgeoning mounds of information that we must deal with on a daily basis in our work and personal lives.

References

[1] S.C. Albright, W.L. Winston, and C. Zappe, *Data Analysis and Decision Making with Microsoft Excel,* second ed. Pacific Grove, Calif.: Thomson Learning, 2003.

[2] R. Amar and J. Stasko, "A Knowledge-Task Based Framework for Design and Evaluation of Information Visualizations," *Proc. InfoVis 2004,* pp. 143-149, 2004.

[3] J. Bertin, *Graphics and Graphic Information Processing.* Berlin: Walter de Gruyter, 1981. Translation of J. Bertin, *La Graphique et le Traitement Graphique de l'Information.* Paris: Flammarion, 1977.

[4] J.S. Brown and P. Duguid, "Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation," *Organizational Learning,* M.D. Cohen and L.S. Sproull, eds., pp. 58-81, Sage Publications, 1991.

[5] S.M. Casner, "A Task-Analytic Approach to the Automated Design of Graphic Presentations," *ACM Trans. Graphics,* vol. 10, no. 2, pp. 111-151, 1991.

[6] T.G. Dietterich, R.H. Lathrop, and T. Lozano-Perez, "Solving the Multiple-Instance Problem with Axis-Parallel Rectangles," *Artificial Intelligence,* vol. 89, pp. 31-71, 1997.

[7] S.G. Eick, "Visual Discovery and Analysis," *IEEE Trans. Visualization and Computer Graphics,* vol. 6, no. 1, pp. 44-58, Jan.-Mar. 2000.

[8] D.H. Freedman, "Is Management Still a Science?" *Harvard Business Rev.,* pp. 26-38, Nov.-Dec. 1992.

[9] D.C. Glaser, R. Tan, J. Canny, and E.Y. Do, "Developing Architectural Lighting Representations," *Proc. InfoVis 2003,* pp. 241-248, 2003.

[10] S. Havre, E. Hetzler, P. Whitney, and L. Nowell, "ThemeRiver: Visualizing Thematic Changes in Large Document Collections," *IEEE Trans. Visualization and Computer Graphics,* vol. 8, no. 1, Jan.-Mar. 2002.

[11] InfoVis 2004 Contest, http://www.cs.umd.edu/hcil/iv04contest/, 2004.

[12] InfoZoom, http://www.humanit.de/en/products_solutions/products/iz/, 2004.

[13] IN-SPIRE, http://www.in-spire.pnl.gov/, 2004.

[14] Inxight Table Lens, http://www.inxight.com/products/oem/table_lens/, 2004.

[15] W. Johnston, "Model Visualization," *Information Visualization in Data Mining and Knowledge Discovery,* U. Fayyad, G. Grinstein, and A. Wierse, eds., pp. 223-228, Morgan Kauffman, 2001.

[16] A. Kobsa, "An Empirical Comparison of Three Commercial Information Visualization Systems," *Proc. InfoVis 2001,* pp. 123-130, 2001.

[17] J. Nielsen and R. Molich, "Heuristic Evaluation of User Interfaces," *Proc. ACM Conf. Computer Human Interaction (CHI 1990),* pp. 249-256, 1990.

[18] S.F. Roth and J. Mattis, "Data Characterization for Intelligent Graphics Presentation," *Proc. ACM Conf. Computer Human Interaction (CHI 1990),* pp. 193-200, 1990.

[19] S.F. Roth and J. Mattis, "Automating the Presentation of Information," *Proc. IEEE Conf, Artificial Intelligence Applications,* pp. 90-97, 1991.

[20] P.M. Senge, *The Fifth Discipline.* Currency, 1994.

[21] J. Sharit, "Allocation of Functions," *Handbook of Human Factors and Ergonomics,* second ed., G. Salvendy, ed., pp. 301-339, New York: Wiley Interscience Publications, 1997.

[22] B. Shneiderman, "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations," *Proc. 1996 IEEE Conf. Visual Languages,* pp. 336-343, 1996.

[23] B. Shneiderman, "Inventing Discovery Tools: Combining Information Visualization with Data Mining," *Information Visualization,* vol. 1, no. 1, pp. 5-12, 2002.

[24] R. Spell, R. Brady, and F. Dietrich, "BARD: A Visualization Tool for Biological Sequence Analysis," *Proc. InfoVis 2003,* pp. 219-225, 2003.

[25] R. Spence and L. Tweedie, "The Attribute Explorer: Information Synthesis via Exploration," *Interacting with Computers,* vol. 11, pp. 137-146, 1998.

[26] R. Spence, *Information Visualization.* ACM Press, 2001.

[27] Spotfire, http://www.spotfire.com, 2004.

[28] E.R. Tufte, *The Visual Display of Quantitative Information,* second ed. Cheshire, Conn.: Graphics Press, 2001.

[29] F. van Ham, "Using Multilevel Call Matrices in Large Software Projects," *Proc. InfoVis 2003,* pp. 227-233, 2003.

[30] S. Wehrend and C. Lewis, "A Problem-Oriented Classification of Visualization Techniques," *Proc. InfoVis 1990,* pp. 139-143, 1990.

[31] J.A. Wise, J.J. Thomas, K. Pennock, D. Lantrip, M. Pottier, A. Schur, and V. Crow, "Visualizing the Non-Visual: Spatial Analysis and Interaction with Information from Text Documents," *Proc. InfoVis 1995,* pp. 51-58, 1995.

[32] M.X. Zhou and S.K. Feiner, "Visual Task Characterization for Automated Visual Discourse Synthesis," *Proc. ACM Conf. Computer Human Interaction (CHI 1998),* pp. 392-399, 1998.

Robert A. Amar received the MS degree in computer science and the BA degree in chemistry from Washington University in St. Louis in 2000. He has been a doctoral student at the Georgia Institute of Technology since 2002. Previously, he has been a consultant for e-business Internet/intranet technologies. His current research interests include task analysis and frameworks for information visualization.

John T. Stasko received the BS degree in mathematics from Bucknell University in 1983 and the ScM and PhD degrees in computer science from Brown University in 1985 and 1989, respectively. He is a professor in the College of Computing and the Graphics, Visualization, and Usability (GVU) Center at the Georgia Institute of Technology. His research is in the area of human-computer interaction with a specific focus on information visualization, peripheral awareness of information, and software agents. He was program chair for the 2003 ACM Symposium on Software Visualization and is papers cochair for the 2005 IEEE Symposium on Information Visualization. He is a member of the IEEE and the IEEE Computer Society.