Richard Brath, London South Bank University, UK; Uncharted Software Inc., Canada
Ebad Banissi, London South Bank University, UK

Using Typography to Expand the Design Space of Data Visualization

Abstract  This article is a systematic exploration and expansion of the data visualization design space focusing on the role of text. A critical analysis of text usage in data visualizations reveals gaps in existing frameworks and practice. A cross-disciplinary review including the fields of typography, cartography, and coding interfaces yields various typographic techniques to encode data into text, and provides scope for an expanded design space. Mapping new attributes back to well understood principles frames the expanded design space and suggests potential areas of application. From ongoing research created with our framework, we show the design, implementation, and evaluation of six new visualization techniques. Finally, a broad evaluation of a number of visualizations, including critiques from several disciplinary experts, reveals opportunities as well as areas of concern, and points towards additional research with our framework.
Introduction

In data visualization, abstract data elements like quantities or categories are encoded into visual attributes of geometry—colors, sizes and shapes—and depicted on an interactive screen or printed on paper. These visual representations act as external memory aids. They facilitate perceptual inferences—spotting outliers, estimating trends and comparing sizes, and enable higher-level tasks such as generating hypotheses and disseminating findings. However, typographic attributes such as bold, italic, and font family variations are rarely used to assign meaning to the data included in visualizations, suggesting a missed opportunity. Using typographic elements to expand the design space of data visualization can enable new types of visualizations, inspire novel applications within existing domains, and lead to potential new areas of economic activity.

This article provides a framework for applying typography and font attributes to data visualizations, and proposes new visualization techniques created using this framework. In the first half of the paper, we perform a systematic review of data visualization theory and practice to identify gaps in the research. We then use cross-disciplinary research to identify existing typographic attributes. These attributes are mapped back to existing research to characterize their use in data visualization, and frame the new expanded design space. In the second half of the article, we review a sample of six new data visualization techniques created using our framework, and consider some early evaluations associated with each. Finally, the results of expert critiques of the broader framework are provided.

The Field of Data Visualization

Data visualization has become a significant area of research in the last 25 years. In the domain of computer science there is a focus on effective visualization and interactive techniques, and the articulation of related data analyses, evaluations, and applications. In visualization, a key step is the transformation of data into a visual representation—called visual encoding. Early researchers, including Bertin and Card, organized the encoding design space into three areas: a) data types, b) visual attributes, and c) marks, as shown in figure 1. This framework is powerful at explaining the construction of visualizations, including some well-known visualization techniques shown in figure 2. For example, the bubble plot encodes quantitative data as x/y location, encodes category data as hue, and renders these using point markers of varying sizes to convey significance. The treemap represents hierarchical quantities with a range of sizes and hues, and renders the data as areas within a whole. The tag cloud encodes word frequencies as size—and in this particular example applies random hues—then renders the words at randomly placed points. The traditional framework continues to inform the teaching and creation of new visualization techniques, as well as formal declarative grammars.
However, if we look at emerging areas of visualization—such as text visualization—there have been fewer innovations. For example, tag clouds are perhaps the most famous text visualization technique to emerge in the last 20 years. Yet they continue to use the very traditional visualization attributes of size and color to adjust words based on data properties. Very rarely do text visualizations venture into the un-researched attributes of font family variations, or **bold** and **italic** scripts. Yet beautiful historic examples using rich typography can be found in other domains. For example, in Carey and Lavoisne’s *A Complete Genealogical, Historical, Chronological, And Geographical Atlas* (figure 3), bold indicates major branches, all caps indicates regions, small caps indicates sovereign rulers, italics represent spouses, and symbols add other information. This effective use of typography suggests a possible gap and opportunity for visualization design.

**What is a Design Space?**

Visualization researchers use the term *design space* without definition. The term is used across many domains, including semiconductors, pharmaceuticals, and human-computer interaction. Some definitions:

- The set of possible designs and design parameters that meet a specific product requirement. Exploring design space means evaluating the various design options possible with a given technology and optimizing with respect to specific constraints like power or cost.  
- The multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality.  
- Design Space Analysis creates an explicit representation of a structured
A design space defines a range of design parameters that can be used to construct possible solutions. It can be a powerful aid, as it frames the exploration of many potential design alternatives, but may also be limiting, as the designer may not search outside the boundaries implied by the design space. Furthermore, designing data visualizations is difficult because there are many trade-offs between design alternatives, so a restricted design space will result in a higher probability of missing a good design. Figure 4, adapted from Munzner, illustrates a design exploration within a design space. First, there are many possible design solutions, some of which are poor, and some better. “The vast majority of the possibilities in the design space will be ineffective for any specific usage context,” explains Munzner. The novice visualization designer (left), unaware of the visualization framework, will be limited to a small portion of possible solutions, while the established designer (middle) typically uses the accepted visualization framework—a broader design space than that of the novice that can yield better results. However, a designer that can use an expanded design space (right) has more potential solutions, including new techniques not feasible within the previous delineations of the design space. To get beyond the existing framework, it is desirable to explore and characterize a broader design space.

**A Method to Expand a Design Space**

Many project-oriented approaches to design do not focus on the design space. For example, user-centered design techniques are focused on user perceptions, behaviors, needs, and experiences. The user-centered approach is focused on the problem space—in other words, characterizing the problem to help direct the design approach and find an optimal solution. While user-centered design may result in a single unique solution that goes beyond an established design space, it does not seek to frame the larger design space.

The goal of this article is to outline the steps of a systematic exploration and expansion of any design space, and walk through that process as applied to the use of text in data visualization. These steps are:

1. Identify gaps in the existing domain’s parameter space, with emphasis on areas that have the greatest potential;
2. Research background across a wide variety of disciplines to identify new parameters, and characterize those parameters both in terms of their originating disciplines and their relationship to well researched parameters in the target domain;
3. Identify novel considerations for new parameters that may impact effectiveness and evaluation;
4. Identify new application areas; then design, implement, and evaluate new kinds of solutions based on these new parameters; and
5. Evaluate the overall results via expert critiques from both the target domain and the originating disciplines.
Identifying Gaps to Pursue in Data Visualization

While data visualization as a research field is more than 25 years old—see early work by Jacques Bertin, William Cleveland, and Jock MacKinlay—there are many gaps and underexplored areas, such as novel visual attributes, encodings, interactions, and evaluations.

At a high level, interactive data visualization transforms data into visual representations perceived and decoded by a viewer. This sequence can be represented as a pipeline, for example, through steps such as data, enrichment, visual encoding, interaction, rendering, viewing, perceiving, and comprehension, as shown in figure 5 (simplified from a diagram by Chen and Floridi).

Each step in the visualization pipeline can enhance the data or introduce unintentional noise and error. As each stage builds on the prior stage, errors accumulate. There will be a gap between reality and perception due to incomplete data, poor design choices, limitations in perception and comprehension, and so on. This gap will always exist, and can be identified at different stages as well as in the overall structure. Therefore, a reexamination of the overall design and each stage of the interactive visualization pipeline can identify gaps. For example, low-level gaps at individual stages may exist due to assumptions in the original concepts. Those assumptions may have been based on existing technical limitations, a narrower scope of uses, or other factors.

Identify Gaps

As opposed to tabular reports, summary statistics, or analytics, the visual encoding step is unique to data visualization. Most visualizations today primarily rely on encoding data into the visual attributes of position, size, and color. These attributes have been well researched. Beyond these attributes, the list of visual attributes can vary considerably depending on the compilation of research. Table 1 shows a compilation of visual attributes as identified by various researchers over the last few decades. Different researchers may group attributes in various ways—the groupings shown here are the authors’.

The final column contains a list of attributes that have been identified by perceptual psychologists as preattentive—attributes that can be automatically perceived regardless of the number of items in a display. Preattentive visual attributes are desirable in data visualization as they can demand attention only when a target is present, can be difficult to ignore, and are virtually unaffected by load. Note how text is included by only a few authors, and how typographic attributes do not appear anywhere on this table.

Some of these attributes—such as texture, shape and text—are rarely used to encode data, and have not been thoroughly researched. There are many possible reasons that some attributes are highly utilized while others less so. For example, size and hue may be popular because they have strong visceral appeal, or because they are easy to code—scale transformations and RGB colors are easily accessible in most programming languages. Another possibility is that they are easily perceived, as shown by preattentive research. Alternatively, computer scientists involved in data visualization may have limited use of other visual attributes, as they have limited knowledge of visual design vocabularies and grammar.
Table 1. Table of Visual Attributes. Visual attributes for encoding data as defined by various information visualization researchers and preattentive vision research up to early 2015 (not including authors’ research).

<table>
<thead>
<tr>
<th>Table of Visual Attributes</th>
<th>Information Visualization Researchersa</th>
<th>Vision Researchb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transform</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position</td>
<td>X X X X X X X X X X</td>
<td></td>
</tr>
<tr>
<td>Length</td>
<td>X X X X X X X X</td>
<td>X</td>
</tr>
<tr>
<td>Size (Area)</td>
<td>X X X X X X X X X</td>
<td>X</td>
</tr>
<tr>
<td>Orientation</td>
<td>X X X X X X X X</td>
<td></td>
</tr>
<tr>
<td>Volume</td>
<td>X X X X</td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape</td>
<td>X X X X X X X X</td>
<td></td>
</tr>
<tr>
<td>Angle</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Curvature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line Ending</td>
<td>X X X</td>
<td></td>
</tr>
<tr>
<td>Closure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edge Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corner Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Icon, glyph, etc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Color</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brightness</td>
<td>X X X X X X X X</td>
<td>X</td>
</tr>
<tr>
<td>Hue</td>
<td>X X X X X X X X X</td>
<td></td>
</tr>
<tr>
<td>Saturation</td>
<td>X X X X X X X X</td>
<td></td>
</tr>
<tr>
<td>Texture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Granularity</td>
<td>X X X X X X X X</td>
<td></td>
</tr>
<tr>
<td>Pattern</td>
<td>X X X X X X X X</td>
<td></td>
</tr>
<tr>
<td>Orientation</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Relation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Connection</td>
<td>X X X X X X X X</td>
<td></td>
</tr>
<tr>
<td>Containment</td>
<td>X X X X X X X X</td>
<td></td>
</tr>
<tr>
<td>Optics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blur</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Transparency</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Stereo Depth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concave/Shade</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Light Direction</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Shadow</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial Occlusion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flicker</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Speed</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Direction</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Miscellaneous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text Labels</td>
<td>X X X X X X X X</td>
<td></td>
</tr>
<tr>
<td>Numerosity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial Grouping</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Artistic Effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arrangement</td>
<td>X X</td>
<td></td>
</tr>
<tr>
<td>Resolution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a See note 14.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b See note 15.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Technological limitations and evolution may also be a driver behind the use of these attributes. For example, the resolution of most displays was limited to 72 pixels per inch until the late 2000’s, limiting the use of such finely detailed attributes as texture, shape, and text. However, much higher resolutions are now
prevalent on mobile devices and retina displays. With fine detail available, each of these attributes may have additional parameters, and hence capabilities that have not been previously explored—for example serifs, italics, and so forth.

Potential Value of Text Visualization
Text is interesting to consider from an economic perspective. As indicated by the definitions of design space, typically the goal is to create value in the context of an objective, such as cost, quality, or profit. There are large amounts of unstructured data—pundits suggest 80% of all data remains unstructured. However, current visualizations of text rarely use font attributes, suggesting an economic opportunity. For example, scimaps.org is a repository of curated exemplars of knowledge maps—data visualizations used to gain insight into the structure and evolution of large-scale information spaces. While 80% of 144 knowledge visualization examples on scimaps.org use text, only a few use font attributes to encode data. Similar results could likely be found on other text visualization platforms like processing.org, d3.js, the Guardian, and Bloomberg. The market for structured data visualization is US$4.1B and growing annually—nearly 10%—with ten major established companies. Text visualization is a nascent market with no dominant text visualization companies, although there exist dominant providers of texts for specific domains—Springer (science), Lexis-Nexis (law), Bloomberg (news), Project Gutenberg (open source books), and so on. If one draws parallels to the market for structured data analysis from relational databases vs. unstructured data analysis from online searches, the unstructured analysis market opportunity emerged many years after the structured market. A similar situation may exist in the opportunity for structured visualization vs. text visualization.

Potential Objections
There are already many existing, peer-reviewed text visualization techniques. The online Text Visualization Browser is a non-exhaustive survey of many peer-reviewed text visualization techniques, with 249 examples logged from 1976-2015, excluding the authors’ contributions (as of Jan 22, 2016). As shown in the left column of table 2, 40 of these have no text, 103 have simple plain text such as axis labels, node labels, document titles or tweet content, and only 106—a mere 43%—use some form of visually encoding additional data into text. When text does encode additional data, the middle list in table 2 shows which combinations of visual attributes were used—size and hue together occur most frequently. In many cases, two or more visual attributes are used to encode data—the list on the right in table 2 shows the count of visual attributes used. Size is used most frequently, closely followed by hue.

Size
Of the 106 visualizations using text encodings with additional attributes, 76 change text size to encode data. Size is highly preattentive—meaning that it can be perceived almost instantaneously. However, using large words reduces the number of words that can be displayed overall, thereby reducing data density. Size variation also interrupts readability for longer passages of text.

Hue
While hue is popular, there can be difficulties using it. Text legibility depends on the contrast between text and the background. For example, it can be difficult to read colored text over varied backgrounds, as shown in the ‘stem and leaf’ plot in figure 6, where red text may occur over an orange bar.
Out of 106 visualizations, we counted 39 tag clouds— in other words, 37% of peer-reviewed text visualizations encoding data into text are variants on tag clouds. But tag clouds have many critics, such as Jakob Nielsen’s: “A one-paragraph summary of each report would probably be more enlightening, be faster to scan, and would take up much less screen space, allowing for more items to be summarized on any given page than tag clouds.”

Font Attributes Are Rarely Used

Only 14 of 249 examples use any kind of font attribute to encode data. In some
cases, the use of these font attributes is fairly simple—for example to indicate a selection highlight. Interestingly, some of these text visualization systems mix and match software components including list boxes, email lists, or search result components, and these other components do use attributes such as bold or underline for example to differentiate a title or indicate a link. This suggests that visualization developers rarely consider these attributes, even when in plain sight! Perhaps the existing design spaces constrain their abilities to notice the opportunities—or perhaps they do not have the requisite design knowledge to apply font attributes.

Cross-Disciplinary Research and Relation to Visualization

To address the gap, a review of type use across other disciplines can be used to find examples of typographic attributes used to encode data. While there is low exposure to typography in the visualization community, other domains, such as typography, cartography, mathematics, chemistry, programming and so on have a rich history with type and font attributes that informs the scope of the parameter space.

Typography

In examples from hundreds of years ago, typographers embedded additional data into texts using font attributes. The table of contents for Chambers’ *Cyclopaedia* uniquely creates a readable paragraph split into a hierarchy enhanced with italics, small caps, roman, and superscript to differentiate between topics, fields, descriptions and chapters (figure 7). Typographers think about text in different contexts—labels, headlines, paragraphs, captions, tables, books and so forth. Type is not homogenous, but has many different applications.

Cartography

Cartographers have long used variation in typography—for example font size to indicate magnitude, as well as font-specific attributes—family, weight, italicization (both forwards and backwards), multiple underline styles (single, double, dashed)
and spacing—as exemplified by Steiler’s Atlas\textsuperscript{26} (top, figure 8). Cartographers differentiate between attributes used to encode category data (such as font family), attributes to encode quantitative data (such as weight), and attributes or combinations thereof (such as case and italics) as seen in a portion of the legend of Steiler’s Atlas in the lower half of figure 8.

**Labeling Across Languages**

Different words, phrases and classifications may need to be compared. Different font families can be used to distinguish which set a particular phrase belongs to. In an example from biology, Haeckel’s Pedigree of Mammals\textsuperscript{27} (figure 9) differentiates across classification systems using roman, italic, blackletter, or a slab-serif typeface.

Similar issues occur with books and prints in multiple languages. Laroon’s book of illustrations The Cryes of the City of London Drawne after the Life\textsuperscript{28} (top left, figure 10) consistently labels 66 etchings in three languages, with English in a roman font, French in an italic font, and Italian in a slightly condensed script font with flourishes. Bloemaert\textsuperscript{29} uses four unique font combinations for four languages, while Bodenehr (after Pyle)\textsuperscript{30} has three. Polyglot phrase books may also differentiate languages, such as de Berlaimont’s Colloquia\textsuperscript{31} (bottom, figure 10), which displays words and phrases from eight languages. Differentiating language by font family can presumably aid cross-referencing across eight columns.

**Other Domains**

Notation systems such as chemical formulas like $[\text{As@Ni}_{12}\text{As}_{20}]^{3-}$, mathematical formulas like $\mu_c(A) = \inf(\lambda^*(O) | O \in \mathcal{F}, A \subset O)$, and markup notations like $<\text{div}$


\textsuperscript{29} Cornelis Bloemaert, “Arion,” Le Temple des Muses (Amsterdam: Chatelain, 1733), accessed April
Figure 9  Haeckel's pedigree chart from 1897 uses different fonts to indicate different classifications. Image courtesy of MBLWHOI Library via archive.org.

Figure 10  Above, illustrations from Laroon, Bloemaert, and Bodenehr (after Pyle) using different font families and caps or condensed fonts to indicate different languages. Below, de Berlaimont’s Colloquia from 1631, a wide phrasebook with 8 languages in 8 columns differentiated with 3 font families. All 3 images above © 2016 by The Trustees of the British Museum; image below courtesy of books.google.com.
class="body">Text</div> use various typographic features such as superscript/subscript, delimiters, and special symbols to add additional information into a line of text. Technical applications – such as surveys, engineering and architecture drawings, software code editors, and alphanumeric charts – use different typographic elements to emphasize, delineate, or otherwise add information to text. For example, there is a long history of alphanumeric financial charts (left, figure 11), and the modern market profile chart (right, figure 11) may use attributes such as case, bold, added marks, and superscripts to add information.32

Software source code presentation relies on a variety of visual attributes to enhance code comprehension. Baecker and Marcus characterized a range of visual attributes – including color, size, italics, small caps, bold, font family – for enhancing source code33 (left, figure 12). These are now commonplace in many modern software code editors. For example, WebStorm (right, figure 12) uses attributes such as background shading, text color, bold, italic, and straight or wavy underline, plus user conventions like CamelCase and programming language syntax requirements of scope (e.g., “.’%.<”) to enhance code readability.

Visualization Domain

While font-attributes are not broadly used in visualization, there are a number of interesting experiments, not necessarily cataloged in the Text Visualization Browser. Some examples are shown in figure 13, including Ellingham’s chart of science (1948)34 that uses size, capitalization, italics, and bold; Skupin’s knowledge
map, with variation in weight and spacing; FatFonts, which vary numeral weights in accordance with magnitude; and TextViewer, with variation in size and multicolor underlines.

Characterization

Based on the cross-disciplinary research, a list of parameters is collected and characterized based on the guidelines from the various fields. Cartographers have differentiated which typographic attributes are relevant to encoding quantitative values (e.g. weight, case) versus categoric values (e.g. font family, spacing). Similarly, Baecker and Marcus provided detailed analysis of each typographic attribute relevant to the formatting of software source code. In typography, one can find fonts where attributes have a range of values. For example, font weight in some typefaces is available in up to 9 weights, and can be used to show quantitative data. Italics are typically forward sloping at 2-20 degrees of inclination, but historically may range from positive to negative, might include vertical italics, and can have slopes at much steeper angles – some surveys from 1881 have an italic slope of 35 degrees. Case is considered more assertive than italics; it allows the designer to stay at the same font size rather than change size (e.g. to indicate magnitude), can be ordered (e.g. as shown in figure 3 from ALL CAPS to Small Caps to Proper Nouns for family branch, sovereign ruler, family member). There are more than 100,000 different fonts, but there are only a few major groupings of font styles into different categories (e.g. sans serif, serif, script, blackletter) although there are a number of different typographic classification systems. Other font-specific attributes include underline, superscript and subscript. Font-specific width includes spacing, condense/expanded, and scaling. There are also paired delimiters like { }, [ ], --, and **, and of course the character glyphs themselves – alphanumerics and symbols. In addition, type designers manipulate other low-level attributes not available to the user of a font. For example, x-height, contrast, stress, and serifs are feasible design parameters, although these attributes are not easily accessible to most applications. All these attributes are shown in figure 14 in the second column. The first column organizes the attributes into 4 groups – the underlying glyphs, attributes made available as part of a font family applicable to all glyphs, attributes that are only applicable when used across a sequence of letters, and attributes typically only available to the font designer.

<table>
<thead>
<tr>
<th>Group</th>
<th>Font Attribute</th>
<th>Visual Channel</th>
<th>Best for encoding:</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glyphs</td>
<td>Alphanumeric Text Glyph</td>
<td>Position</td>
<td>D</td>
<td>ape bat cat dog 123 456</td>
</tr>
<tr>
<td></td>
<td>Symbols</td>
<td>Length/Size</td>
<td>D</td>
<td>! ? @ comment $ var</td>
</tr>
<tr>
<td>Font Family Attributes</td>
<td>Oblique / Italic</td>
<td>Intensity</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Case/Normal/Caps</td>
<td>Orientation</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Typeface</td>
<td>Change</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Underline</td>
<td>Pre-attentive</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Condensed</td>
<td>Potential</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Squished</td>
<td>Font/Color</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spacing</td>
<td>Font/Color</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Super/subscript</td>
<td>Font/Color</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delimiters</td>
<td>Font/Color</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Font Design</td>
<td>X-height</td>
<td>Font/Color</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Contrast / Stress angle</td>
<td>Font/Color</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Serif length / Bracket size</td>
<td>Font/Color</td>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>

* / + indicates primary / secondary visual channel for font attribute
† HP: Highly probable, P: probable, D: doubtful
Note: does not include traditional visualization attributes such as size, color, texture, outline, blur, shadow, etc., which are applicable to any marker including type.

Figure 14 Overview of font-specific visual attributes; related visual channel, preattentive potential, best encoding and example. Image © 2016 by Author.

35 André Skupin, In Terms of Geography, courtesy of André Skupin, San Diego State University, “1st Iteration (2005): The Power of Maps,” in Places & Spaces:

36 FatFonts, which vary numeral weights in accordance with magnitude; and TextViewer, with variation in size and multicolor underlines.

37 with variation in weight and spacing; FatFonts, which vary numeral weights in accordance with magnitude; and TextViewer, with variation in size and multicolor underlines.

38 Similarly, Baecker and Marcus provided detailed analysis of each typographic attribute relevant to the formatting of software source code. In typography, one can find fonts where attributes have a range of values. For example, font weight in some typefaces is available in up to 9 weights, and can be used to show quantitative data. Italics are typically forward sloping at 2-20 degrees of inclination, but historically may range from positive to negative, might include vertical italics, and can have slopes at much steeper angles – some surveys from 1881 have an italic slope of 35 degrees.

39 Case is considered more assertive than italics; it allows the designer to stay at the same font size rather than change size (e.g. to indicate magnitude), can be ordered (e.g. as shown in figure 3 from ALL CAPS to SMALL CAPS to Proper Nouns for family branch, sovereign ruler, family member). There are more than 100,000 different fonts, but there are only a few major groupings of font styles into different categories (e.g. sans serif, serif, script, blackletter) although there are a number of different typographic classification systems. Other font-specific attributes include underline, superscript and subscript. Font-specific width includes spacing, condense/expanded, and scaling. There are also paired delimiters like { }, [ ], --, and **, and of course the character glyphs themselves – alphanumerics and symbols. In addition, type designers manipulate other low-level attributes not available to the user of a font. For example, x-height, contrast, stress, and serifs are feasible design parameters, although these attributes are not easily accessible to most applications. All these attributes are shown in figure 14 in the second column. The first column organizes the attributes into 4 groups – the underlying glyphs, attributes made available as part of a font family applicable to all glyphs, attributes that are only applicable when used across a sequence of letters, and attributes typically only available to the font designer.

Using Typography to Expand the Design Space of Data Visualization 71
Relation to Data Visualization

To further characterize font attributes, they can be mapped to related research used by data visualization researchers. Early on, Gestalt psychologists recognized that similar elements will be perceived as part of a group. Perceptual psychologists and vision research further identify different visual channels such as position, size, intensity, orientation and shape. Each font attribute utilizes some combination of these perceptible visual channels, as shown in the middle columns of figure 14. For example, font weight at a macro-level can be understood to primarily alter the intensity of characters, which it achieves at a micro-level by varying the widths of letter strokes.

Some visual channels are stronger cues for guiding attention than others, and can visually pop-out. This is summarized in the column in figure 14 labeled “preattentive potential.” For example, intensity or luminance is considered probably preattentive, and line width or size is considered undoubtedly preattentive. Font weight—which uses both—is recorded as highly probable in this table.

Some visual attributes are more accurate for perceiving magnitude, and the related concept of the number of distinct levels that can be perceived. For example, some channels are effective for encoding quantitative data (e.g. size), while others can only differentiate (e.g. shape). Font weight, using size and intensity, can be used for encoding quantitative or ordered data, while font family, using shape, can be used for indicating categories. This is shown in the second last column of figure 14. In most cases, font attributes do not have many distinct levels, limiting their use to only a few different data levels.

Glyphs – α, β, γ, δ, ε, ζ, η, θ, χ, ω – can be used to encode literal data. Alphabetic or numeric glyphs can also be used to order data, based on a learned ordering. Both literal encoding and ordered encoding are not constrained to a few values like other attributes; however, these are highly unlikely to be pre-attentive.

The final column in figure 14 provides an example of the font attribute, including variations of the attribute across several levels. For example, variations in underlines can be used to create an ordering as shown in the table (and in figure 8).

Unique Considerations: Legibility and Readability

A new palette of tools gives rise to a new set of considerations. As revealed in cross-disciplinary research, legibility and readability are of paramount concern to typographers and cartographers. Legibility is a perception issue. It is concerned with the ability to clearly decipher individual characters as well as commonalities within a font that increase letter identification. Typographers and psychologists discuss factors at the level of characters—whether these have consistent stroke widths, open counters, or wider proportions; between characters— their risk of error, run-together risk, x-height; across a series of letters—including font-tuning, or predictability across letters; and environmental factors such as illumination and distance. Readability, on the other hand, is a comprehension issue concerned with the ease of reading lines and paragraphs of text, and can also be affected by many factors such as line length, kerning, leading, x-height, and font weight. Visualization researchers are unfamiliar with these concepts—a recent study on text highlighting did not consider legibility.

Applications in a New Design Space

Simply outlining the parameters of a design space does not provide any indication of how any new capabilities might be used. How can they generate new value? 

Value is unlikely to be uncovered by simply converting an existing successful technique to a new parameter. For example, simply changing a tag cloud to use font
weight instead of font size is perhaps more space efficient, but would be viscerally less appealing and would not solve any new problems (figure 15).

Expanding the design space using cross-disciplinary research should lead to the development of many demonstrable applications. Typographers structure type design into a type hierarchy ranging from the scope of glyph design to words, sentences, paragraphs and documents, to systems applied across many documents—for example, a design system used across a series of books, or guidelines for corporate branding and visual identity. This hierarchy is somewhat similar to the differentiation between representations of marks as point, line, and area as discussed in data visualization. Also, cartographers differentiate between encoding quantitative data like font weight vs. categoric data like typeface, in addition to the literal encoding of the text itself, which is similar to the visualization classification of data types. Combining these creates the $3 \times 6$ design space of possible applications depicted in figure 16. Typographic scope is listed vertically, data types are listed horizontally, and cell intersections identify applications. For example, cell QW indicates embedding Quantities into Words, which could be achieved with varying font weight or italic slope based on data as shown in the sample column; or other novel approaches such as varying the length of a word's underline to indicate a quantity. Cell CG indicates embedding Categoric data into Glyphs—imagine these as a subset of a word—for example, indicating silent letters with a lightweight font, as shown in the example Gloucester.

Figure 16 can be combined with the 15 font-specific attributes outlined in figure 14 as an additional dimension. Combined, they create a large $3 \times 6 \times 15$ design space, which suggests the possibility for dozens of techniques, if not hundreds. Figure 14 also hints at other types of data encodings—grouping—which could also be pursued outside the current scope of this investigation.

This framework can be used to review the existing text visualization literature. For example, the scope of textual encodings used in the examples at the Text Visualization Browser is outlined in table 3. 40 out of 249 visualizations do not use text. Out of the remaining 209, 174 visualizations operate at the level of words—that is, for 83% of the examples, the scope of the text is words. These may be LW (Literal Words) such as labels on a scatterplot; CW (Categoric Words) such as color-coded

Figure 15 (Above) Tag cloud using size (left) or using font weight (right) to encode the revenues of Fortune 100 companies. Image © 2016 by Author.

Figure 16 (Below) Application opportunities for font attributes—typographic scope vs. data type. Image © 2016 by Author.
Table 3. Scope of text used in text visualizations in Text Visualization Browser.

<table>
<thead>
<tr>
<th>Scope of Text in Visualization</th>
<th>None</th>
<th>Literal Only</th>
<th>Categoric Encoding</th>
<th>Quantitative Encoding</th>
<th>Total Categoric Encoding</th>
<th>Font-Specific Encodings</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyph</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word</td>
<td>87</td>
<td>19</td>
<td>68</td>
<td>174</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line</td>
<td>9</td>
<td>3</td>
<td>7</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paragraph</td>
<td>8</td>
<td>4</td>
<td>3</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Document</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corpus</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

words; or QW (Quantitative Words) such as variably-sized words. Beyond words, larger scopes of text encoding data occur infrequently. Sentences—a title, a tweet or a keyword in context—occur 19 times. Paragraph scope—an abstract or a few sentences—occurs 15 times. No examples exist depicting text at the level of a corpus; current examples of corpus visualization reduce documents down to marks such as a dot per document, or perhaps lists of words describing a topic. In general, word-based visualizations of text may be common because larger text sources can be reduced down to word lists using analytic techniques such as word frequency, entity extraction, topic classification and so on. At the sub-word level, only two text visualizations are listed—one for the analysis of suffixes, and the other phonetic units.

With regard to font-specific attributes like bold, italic, and underline, there are very few visualizations, as shown in the last two columns on the right in table 3, and are all categoric encodings—there are no font-specific quantitative encodings.

While table 3 does indicate that there is some exploration in scope and encodings, there is clearly a strong bias for words. Furthermore, as indicated in table 2, the existing use cases are almost entirely biased to use of size and/or hue to encode data—and as discussed earlier, there are various issues with these visual attributes, including readability, information density, and contrast. Also, as indicated by table 3, there is very little exploration within typographic attributes. Given a 3 × 6 × 15 design space, there are only 16 instances. There are many areas that remain entirely unexplored.

The next step is to identify new and potentially valuable applications for which new visualizations using font attributes can be constructed and evaluated. The application of text and font attributes to visualizations does not need to be constrained to common typographic nor visualization conventions. Font attributes can be applied to subsets of words, sentences can flow along paths, internal properties of glyphs can be modified by x-height or by font width. Given the space limitations of the article, a sampling of six unique problems and designs created through this investigation will be shown, each addressing a different area of the data type x scope matrix—LS, CD, QW, QP, QS and QG. An evaluation of the particular technique within each example will be followed by critiques of the larger body of work afterwards. As this is a work in progress, note that evaluation and critiques are ongoing—refinements and additional tests for some techniques as well as new techniques are still being generated in relation to the design space.

**LS: Literal Sentences — microtext lines**

Let us begin with simple example. Line charts are frequently used for timeseries analysis. A simple line chart with one line doesn’t require labeling—the title can unambiguously indicate the line. But when a few lines are added, cross-referencing between a label and a legend requires some cognitive effort. One of the benefits
of diagrammatic representations is reduced cross-referencing—for example, the need to refer back and forth between a line in a chart and a legend with a description. Furthermore, for some types of data, legends can be quite verbose. Figure 17 plots retweets over time for the most popular Donald Trump tweets beginning in late August 2015. Instead of a separate legend, tweet content replaces the line.

By replacing the line with the content, the viewer’s next question—what did he say?—can be immediately addressed in context. In this example, the orange tweet was quickly retweeted due to the huge fan-base of one @AshtonSSOS—Ashton Irwin, the drummer for the Aussie teen band 5 Seconds of Summer—who wrote the initial tweet, whereas the slower blue line likely grew in popularity due to the comedic content.
This approach can scale to many more lines. Figure 18 shows the same approach, but displaying more than 35 timeseries of country economic data. Each line is labeled with microtext indicating the unique country name in multiple languages as well as redundantly encoded with unique colors and unique typefaces/weights.

A small group of six expert users were provided with variations of the chart in figure 18, including a plain line chart with 35 lines and endpoint labels, and a line chart with 35 lines displayed as microtext. The users were from the domain of financial services—they use line charts every day, all day. After seeing the traditional chart with only solid lines and end point labels, 4 out of 6 experts did not answer a simple question—Which country had the highest unemployment in 2002? When provided with lines directly encoded as text (as in figure 18), the experts immediately identified individual lines, and could also answer more complex questions requiring assessment of the trend of a line over time in comparison to other lines—How did Switzerland fare through the financial crisis of 2008? While 5 out of 6 experts were positively disposed to this technique, one user could not discern the small, six point font used in the paper examples provided. Interactive control over font size could have remedied this issue.

**CD: Categoric Document—typographic mosaic**

Some documents are big tables, or lists. Many visualization techniques are used to count things from lists. Distributions, treemaps, mosaic plots, and sometimes pie charts and bar charts may represent numbers of things by area. However, these visualizations depict only a summary—the underlying elements have disappeared. For example, the viewer’s task may be to locate where a particular item occurs, rather than any overall tallies, or they may be interested in items adjacent to a particular target, and so on. With only sums, getting to the underlying data requires additional interaction, such as a tooltips or mouse clicks. Interactions are slow when compared to simply shifting attention, and thus it becomes desirable to directly depict the items that make up the sum.

The *Titanic* dataset is a popular sample dataset for data visualization. There are 1308 passengers, all of which can be categorized by age, gender, class and survivorship. Visualizations of the *Titanic* data typically reduce the data down to summaries, then plot the sets, for example, as a mosaic, a Venn diagram, Parallel Sets, and so forth. However, Alsallakh et al.’s report on state-of-the-art set visualization identifies 26 analytic tasks related to use of sets, of which 12 pertain to underlying set elements, the elements’ attributes, or relations between elements. Unfortunately, summary techniques do not retain the individual elements that are needed for almost half of the tasks.

With high-resolution displays, thousands of individual items can be explicitly labeled and the overall area formed by a group of labels indicates the counts of the individual items. Figure 19 displays all 1308 passengers on the *Titanic*, grouped vertically by class (1, 2, 3). Within each horizontal band, both type and color are used to distinguish between men (plain/above) and women & children (italic/below), and survivorship—died (red serif/left) and survived (green sans serif/right).

Macro-patterns are visible, such as the higher proportion of survivors in first class relative to other classes. At the same time, the detailed names of each individual are immediately accessible. Similar to the Vietnam Veterans Memorial in Washington, D.C., each person is made visible. Macro-questions can be asked of this graphic—*Were women and children really first across classes?*—as well as micro-questions—Did the Astors survive or die?

Using labels to form areas that can be perceived and compared as quantities assumes labels are of similar length within each area. Unfortunately, surviving first class women’s & children’s names average 36.7 characters in length while deceased
third class men average only 22.0 characters. This is remedied in this example by shortening all names to a familiar name and a surname. For example, the passenger recorded as Brown, Mrs. Thomas William Solomon (Elizabeth Catherine Ford) is reduced in the visualization to Elizabeth Brown. Using these reduced passenger names, the same two segments average 13.9 and 13.6 characters respectively—a 2% difference, well within the margin of error of area perception.

**QW: Quantitative Words—labeled cartograms**

Reviewing cartography yields both typographically rich maps, such as the earlier example in figure 8, as well as maps devoid of labels, such as choropleth maps (see figure 20). Choropleth maps fill regions with different colors to indicate data values—they have existed for almost two centuries, and they are extremely popular. However, choropleth maps have many well-known problems:

- Regions with large areas are much more visually salient than those with small areas
- Some small areas may not be visible at all—like Singapore or Luxembourg on a world map.
- Not all viewers are familiar with geographic shapes—63% of young Americans could not locate Iraq on a map of the Middle East in a National Geographic survey in 2006.
- It can be difficult to depict additional data attributes on the map, although it can be achieved with techniques such as added glyphs per country, or textures like stripes.

A label-based cartogram can be used instead of country shapes. If ISO 3166 codes are used, then all labels will be a consistent length (e.g. 3 letter codes). Labels can be placed such that each label is clearly visible and retains local proximity to adjacent countries. Furthermore, while a choropleth map typically depicts only one numeric value via color, a label can depict multiple attributes. Typical visual attributes such as size and color can be used to represent data. Alternatively, equal
sizes can be used for labels with quantities encoded as font attributes, as shown in figure 21.

A small study was conducted to compare the choropleth map and the labeled cartogram for identification and location tasks with two groups—ten undergraduate university students and seven information visualization professionals. Comparable maps were created with a single variable and consistent color encoding of data. Tasks were similar to the National Geographic 2006 geographic literacy survey. The identification task marked a particular country on the map and required the viewer to name the country. The location task required the viewer to find a specific country on the map and report the color of that country. Countries used in the tasks were not familiar ones, that is, not highly populated, not part of the G20 nor Western Europe, nor frequently in the news. Each viewer had a set of eight questions evenly distributed between the two task types and the two map types.

The labeled cartogram outperforms the unlabeled choropleth map in both tasks as shown in Table 4. For the identification task, ISO labels significantly
outperform the choropleth with 65% correct answers vs. 15% correct. This may be due to the mnemonic nature of ISO country codes—on a choropleth map, an arrow may be pointing to a shape with few mnemonic affordances, whereas an arrow pointing at a mnemonic code such as SLE, may trigger recognition of Sierra LEone. This small study should be repeated with a larger group of subjects to determine whether these differences are similar in a larger population.

**QP: Quantitative Paragraphs – skim formatting**

Skimming is a reading technique where the eyes rapidly sweep across a large body of text to spot the main ideas and gain an overview of the content. At a low level, the strategy requires the reader to dip into the text looking for words such as proper nouns, unusual words, enumerations, etc. To make uncommon words pop out, first each word in the document can be tagged by its usage frequency in the broad language. Then each word can be assigned a different font weight such that the least frequent words have the heaviest weight down to the most frequent words that have the lightest weight. Figure 22 shows the opening paragraph of *A Tale of Two Cities* formatted to facilitate skimming. Words such as wisdom, foolishness, belief and incredulity pop out.

Skim formatting has received feedback from multiple potential user communities with highly positive responses such as “I can see using this immediately in my own visualization research,” or “The technique can work well by aiding recognition of keywords instead of relying on searching (recall).” Additional improvements could be made by considering the application of word recognition research. For example, if short words and function words are frequently skipped, could their emphasis be further reduced? Another improvement would be to reconsider which content is bolded – if the objective is comprehension of main ideas then advanced machine learning algorithms could potentially be used to identify the most salient portions of the text to be most heavily weighted, rather than inverse word frequency.

**QS: Quantitative Sentences – proportional encoding**

There are various applications where short sentences appear in user interfaces such as news, online searches, and social media. These are often used for document titles, news headlines, email subject headings, tweets, search results (keyword in context), pull-quotes, and so on. As such, there may be much more associated information – for example, dates, authors, sources, subjects, and document properties such as...
document length and number of readers. Interfaces that list these sentences often don’t list other metadata, or use additional columns to itemize related information in a textual format that does not stand out. In some cases, traditional visual attributes such as size and brightness may be used to encode quantities. For example, NewsMap.jp is a treemap of headlines where headline size indicates the number of related articles, and brightness indicates recency.68

Using font-based techniques, quantitative metadata about documents can be depicted in addition to their titles. Figure 23 shows a list of Today’s Featured Articles from Wikipedia. Each line shows the title and a portion of the initial sentence of an article. The length of a particular format indicates a quantitative measure. For example, the length of bold is an indicator of the article length—articles regarding Richard Nixon and Barack Obama are particularly long, while Action of 1 January 1800 is short. Underline indicates readership—the most read article is The Green Children of Woolpit. Note how some formats—like bold length—are easily perceived, whereas others require more focused attention. This was predicted by the preattentive potential previously discussed (central column, figure 14).

This approach can be extended in a few ways. The quantitative data attributes can be used to sort the lines, and relative areas of a particular format can be compared. Figure 24 shows two sets of movie reviewers’ comments from the websiterottentomatoes.com, where the length of bold is used to indicate each reviewer’s score. In this case, the original reviewer’s commentary was truncated to a set length, and padded if the reviewer’s quote was extremely short. Scores were normalized to a common range of 1–10. As the number of reviews varies per movie, the reviews were sorted based on score, then sampled at regular intervals across the full range of reviews to extract a common number of reviews. Plotting these reviews with the length encoding score reveals patterns. At a macro-level, the movie Toy Story 3 can be

---

**Figure 23 (Above)** List of articles from Wikipedia with length of format indicating article length (bold length), page views (underline), newness (case), and number of authors (italic). Image © 2016 by Author.

**Figure 24 (Below)** Movie reviews for Toy Story 3 and Frozen with length of bold indicating score. Toy Story 3 has more bold indicating a higher overall score. Reviews are sorted by score, Frozen has a lower slope indicating higher dispersion. Image © 2016 by Author.

---

seen to have more dark bold, indicating higher review scores than the movie Frozen.
The slope is an indication of dispersion—Frozen has a wider range of scores, with one reviewer providing a particularly poor review and thus almost no bold. And, at a micro-level, the text is immediately accessible.

Proportionally encoded lines of text are spatially and perceptually efficient. A set of headlines and a single associated quantitative value were compared across three different encodings: 1) size—as in a treemap, 2) font weight—with 5 different weights applied evenly across the entire sentence to indicate 5 different data values, and 3) proportion (figure 25). The treemap uses size within a fixed area, making some items large, and thereby decreasing the amount of space available to other items—with some headlines too small to read. By contrast, encoding data into font attributes allows a fixed text size to be used, making all headlines readable. The three different encodings of news headlines were compared using different news data sets. The proportionally encoded headlines consistently outperformed the other two representations, with more readable headlines and overall lower information lossiness.

In terms of perceptual efficiency, comparison of line lengths—used by proportional encoding—has the lowest margin of error (+/−2.5%). Comparisons of rectangular areas—used by treemaps—have a higher margin of error (+/−5%). Font weight used by itself is the least efficient, as it only offers a few levels that can be readily perceived.

**QG: Quantitative Glyphs—prosodic text**

Song in prose is often minimally differentiated from other text, for example, by being set in italics. However, this does not convey any of the prosodic song qualities such as note pitch or duration. While traditional music notation could be used, this would interrupt the flow of the text and require a lot of space. Instead, each syllable can be encoded independently with low-level typographic attributes. In figure 26, lower-case fonts have been compressed//expanded to indicate note duration; and the lower-case font x-height and baseline have been shifted to indicate note pitch—in other words, tall letters indicate deep pitches, while short letters indicate high pitches.

This visualization has not been evaluated. There are many potential issues still to be considered. For example, x-height cannot represent a wide range of pitches, changes in x-height and font-width can be disruptive to reading, fonts with high aspect ratios are less legible, and so forth.

**Evaluation**

Each of the individual visualizations can be evaluated for specific goals as shown in the preceding examples. Techniques included measurement of information density on proportionally encoded lines of text (QS); measured encoding accuracy on the

---

**References**


70 Heer and Bostock, “Crowdsourcing Graphical Perception,” 203–12.
A significant portion of the examples have been critiqued by typographers. A subset of examples similar to those shown above were presented at a type conference,74 followed by reviews with specific attendees afterwards, including at least one online review.75 Reviews have generally been positive, along with appropriate skepticism for some techniques. Interestingly, typographers raise different issues than visualization researchers.

- **Legibility** is a key concern to typographers, but one unexpressed by visualization researchers. Legibility was typically not an issue in the visualizations, except where too many levels of particular font attribute were not perceivable—for example, the skim formatting of Dickens text in figure 22 utilizes five levels of font weight, but not all levels are uniquely perceivable.

- **Readability** is another important concern for typographers. In particular, a number of typographers expressed concern that changing more than a singular attribute can impact readability. For example, figure 23 simultaneously varies weight, underline, italic, and case—which reduces readability. Opinions vary—for example, Kane indicates multiple attributes can be combined to create contrast between words,76 whereas Binghurst indicates that a sudden shift across multiple type attributes does not follow conventional typographic grammar.77 Further typographic examples that mix many attributes can be found. For example, dictionary entries use text with many simultaneous variations in font attributes (figure 27).78 Tracy suggests that readability is not as important in works such as directories or tables, where the viewer is not reading continuously but searching for a single item of information.79

If the primary task is reading, then even a singular strong cue such as font weight may be disruptive, because it is difficult to ignore. Some typographers consider italics as form of quiet emphasis, less disruptive to reading than a strong form of emphasis such as bold.

- **Interactivity** is a potential means to address issues of readability and also enhance functionality. One typographer demonstrated that the skim formatting technique could be easily toggled back and forth between the non-formatted text and the skim-formatted text. In skim formatting, some words are lighter and slightly narrower, while some words are heavier and slightly wider. On the balance, the overall line-lengths remain the same, preserving...
Many, on opening the Encyclopedia of needlework will be disposed to exclaim as they read the heading of this first section: What is the use of describing all the old well-known stitches, when machines have so nearly superseded the slower process of hand-sewing? To this our reply is that, of all kinds of needlework, Plain Sewing needs to be most thoroughly learned, as being the foundation of all. Those who are able to employ others to work for them, should at least know how to distinguish good work from bad, and those who are in less fortunate circumstances, have to be taught how to work for themselves.

Many, on opening the Encyclopedia of needlework will be disposed to exclaim as they read the heading of this first section: What is the use of describing all the old well-known stitches, when machines have so nearly superseded the slower process of hand-sewing? To this our reply is that, of all kinds of needlework, Plain Sewing needs to be most thoroughly learned, as being the foundation of all. Those who are able to employ others to work for them, should at least know how to distinguish good work from bad, and those who are in less fortunate circumstances, have to be taught how to work for themselves.

the relative positions of words across both formats (see the side by side comparison in figure 28). This allows for an interactive toggling between both formats, and the eye can maintain focus on a particular position in the text across the transition. Thus, interactive toggling can be a means to achieve the most effective readability and skimming. Interactive font-attribute manipulation had not been considered prior to this.

• **Intuitive Mappings** make it easy to decode multi-attribute labels. Most of the examples use encodings that have been specifically chosen, as opposed to randomly assigned. For example, figure 21 shows heavier font weights which map intuitively to larger expenditures. This in turn allows for multi-variante labels to be created that are engaging and stimulating, rather than cognitively burdensome, although there is a tradeoff with the readability issue for multi-variante labels.

• **Semantic Encoding** is an issue raised by both typographers and visualization researchers. Font families are often described with different adjectives and properties. For example, Comic Sans is for informal communications, black letter may be associated with the medieval period or heavy metal bands. In The Destruction of Syntax (1913), futurist poet Marinetti says “On the same page, therefore, we will use three or four colors of ink, or even twenty different typefaces if necessary. For example: italics for a series of similar or swift sensations, boldface for the violent onomatopoeias, and so on. With this typographical revolution and this multicolored variety in the letters I mean to redouble the expressive force of words.” Apollinaire uses typography and layout in the poems in his *Calligrammes* to convey meaning as well as words (figure 29).

---

79 Tracy, Letters of Credit, 31.
Similarly, comics and manga have developed typographic conventions to convey semantic information in text bubbles, and overlays for sound effects—small letters for whispering, drippy letters for sarcasm, rising baseline for rising voice, squiggly text for gasping, and so on (figure 30). None of the visualization examples provided consider the semantics of font and how those might be encoded, nor does the framework as provided have any indication of how and where the semantic elements of typography fit. This is an area suitable for future work.

Information Visualization Critiques

This font-attribute visualization design space has been critiqued by five InfoVis experts on two separate occasions, as well as given more general feedback. There seems to be interest at two different levels—individual words, since labels are pervasive throughout data visualization; and everything else—lines, paragraphs, documents, glyphs. Criticisms and discussion include:

- **Is this even data visualization**: The biggest contention is whether or not text can be considered an element of visualization. Purists may define visualization as a process that utilizes visual encodings that are preattentive—automatically perceived—while text requires active attention to be perceived—controlled processing. Therefore, text is not visualization. However, font-attributes such as **bold** clearly utilize preattentive features, thus this argument is really focused on whether the framing of the design space includes literal encodings. Furthermore, we argue that literal encodings are actually part of the design space. For example, replacing a dot on a scatterplot with an alphanumeric character preserves the initial reading of the scatterplot, increases data density by adding additional information with the character, and allows for perception of associative micro-patterns—local adjacencies—that would otherwise require interactive techniques to reveal.

- **Are these really new attributes**: One researcher considered font-attributes just variants on established attributes. For example, font weight is effectively the same as intensity, since both representations vary the amount of “ink” encoding the data. From a typographic perspective, these are not the same—font weight varies the stroke width, maintaining high contrast against a background regardless of the weight. Intensity, however, varies the brightness of the text, thereby reducing contrast to the background and reducing legibility. At the end of 2015, 9 examples in the Text Visualization Browser used intensity to encode data, while none used weight—suggesting a lack of awareness that a similar effect to brightness could be achieved using font weight while enhancing contrast and legibility.

- **Aesthetics**: Some researchers expressed personal opinions that font attributes
do not have the same visceral appeal as bright colors or size differential. This criticism indicates both the strength and weakness of type. Type is unlikely to replace other existing visualization techniques when dramatic visual representations are desired. The corollary is that type may be particularly effective for other tasks such as analytical tasks or monitoring tasks.

- **Label Length Problem**: Using labels instead of markers such as dots raises a potential issue where longer labels may be more salient than shorter labels. For example, in the scatterplot of national parks in figure 31, the length of park names varies widely, from the lengthy *Great Smoky Mountains* to the brief *Zion*. The concern is that long labels have greater prominence and can bias perception, particularly if long labels are concentrated in one area. There are many possible solutions. Cartographers often do not make adjustments—*Peru* and *Philippines* remain the same length. The equal area cartogram uses consistent three letter country codes, as shown in figure 21. *Peru* becomes PER and *Philippines* becomes PHL. In the case of the *Titanic* passengers (figure 19) all names are contracted to simply given name + surname. Size could potentially be used, but could be confused with the convention commonly found on maps of using size to encode data. String lengths could be normalized, as some fonts come in variable widths, and intercharacter spacing can be adjusted—the combination of these two could be used to compress long names and stretch short names. Finally, interactions to toggle between alternate marks such as text and dots could be used. Future work could include a set of experiments to evaluate whether there is a perception problem—perhaps people are conditioned by familiarity with map-based encodings, and would not be as prone to errors as anticipated. Assuming that problems do exist, various alternative solutions should be evaluated.

- **Multi-attribute Interference Problem**: While there are many typographic attributes that can be used, using many at the same time in the same text can interfere with the perception of each attribute. For example, in the proportionally encoded Wikipedia headlines (figure 23), the length of uppercase is difficult to perceive mixed with all the other changes in italics, bold, and

---

Figure 31  Label length varies widely: long labels are more dominant. Image © 2016 by Author.
underline. Case alone may be easy to perceive. Understanding how these different attributes can be used together is important. Ware suggests that different visual channels should be used, and these should be perceptually separable as opposed to integral. Ware, Information Visualization, 157–72.


86 Ibid.

Economic and Innovation Discussion
Critiques suggest that there is some value in these approaches, and that they can be applied at different levels—from characters, to words, to longer text. Whether or not these techniques have economic value is an open question. One measure is perhaps the breadth of interest in a new technique. Research in font-specific attributes applied to visualization has gained interest from different domains including text visualization, information visualization, typography, and information search, including the author’s research, as well as that of Strobelt.

Moreover, these frameworks define areas for future innovation that remain untouched. For example, document corpuses when analyzed are reduced to topics, keywords, graphs, and such. But the design space suggests the possibility of textual views across an entire corpus—which is now technically feasible using extremely high-resolution display walls such as Hyperwall2, a 250 mega-pixel display. Or font-attributes such as paired delimiters may have some novel encoding potential for indicating groupings among non-alphanumeric attributes such as glyphs or table cells.

Finally, the cross-disciplinary context implies that, at a minimum, the use of text in visualization can be aesthetically enhanced, given the beautiful examples that exist in high-resolution print environments such as maps and genealogical charts. Research into aesthetics can be challenging to perform in the domain of evidence-based, experimental methods used in computer science and new methods for experimenting with aesthetics may be required.

Other Questions
The analysis and discussion raises interesting questions about evaluation of the broader visualization pipeline. Optimizing a design for one stage of the pipeline may impact other stages. For example, an encoding that can be perceived quickly does not mean it can be decoded easily. In the labeled cartogram (figure 21), searching for textual encoding of countries may perhaps be slower, but the decoding of countries may be enhanced by mnemonic three letter codes which facilitate recognition. Typographers point to Tall Man Lettering as an example where type is designed to perceptibly make some syllables stand out and be specifically disruptive to reading. This is used to differentiate key syllables in look-alike drug names, which appear on computer screens, on tiny labels, at dispensaries, and in

Sample Tall Man Lettering of Look-Alike Drugs

<table>
<thead>
<tr>
<th>Uppercase</th>
<th>Lowercase</th>
<th>Tall Man</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDROXYPROGESTERONE</td>
<td>medroxyprogesterone</td>
<td>medroxyPROGESTEROne</td>
</tr>
<tr>
<td>METHYLprednisolone</td>
<td>methylprednisolone</td>
<td>methylPREDnisolone</td>
</tr>
<tr>
<td>METHYLtestosterone</td>
<td>methylestosterone</td>
<td>methylTESTosterone</td>
</tr>
<tr>
<td>VINCristine</td>
<td>vincristine</td>
<td>vinCristine</td>
</tr>
<tr>
<td>HYDROXYZine</td>
<td>hydroxyzine</td>
<td>hydrOXYzine</td>
</tr>
<tr>
<td>HYDRAZINE</td>
<td>hydrazine</td>
<td>hydrAzone</td>
</tr>
</tbody>
</table>

Figure 21 Tall Man letters accentuate differences in central syllables of similar drug names to reduce potential for error. Image © 2016 by Author.
other applications (figure 32). The goal is to reduce errors of confusion, which can be fatal. The design solution is a counter-intuitive disruption of the differentiating syllable.

More holistic evaluations aligned with broader system goals need to be considered. This has long been an issue discussed between typographers and experimenters. For example, Dillon points out that ergonomists are so concerned with control over variables that the experimental task bears little resemblance to the activities that most people routinely do when reading. Similarly, some visualization designers point to the need for broader evaluative techniques, such as Tory and Möller and Kosara.

**Conclusion**

At the broad level of expanding any design space, this particular investigation has shown how a systematic investigation into a design space can be used to re-frame the design space and create new applications.

The method shows how interdisciplinary research can be used at multiple points in the process. Early on, cross-disciplinary reviews of background research were used to reveal new parameters. Then a review of the framing of the design space in other disciplines helped frame the expanded space, and also suggested possible applications like skimming and prosody. Finally, cross-disciplinary expert critiques exposed specific issues that may not be known to other domains, identified aspects missed in the evaluation of individual techniques, and uncovered broader issues in holistic evaluation. Cross-disciplinary expert critiques led to a more robust definition of the design space and evaluation criteria, such as the introducing legibility and readability into data visualization.

The characterization of font attributes, the simple evaluations of the individual techniques, and the insights gained from the critiques suggest that there is much more work to be done via evaluation. As this is ongoing research, more in-depth evaluation can be carried out for the six specific techniques shown. There are also higher-level questions that span the design space. For example, given multiple simultaneous font attributes, what is the tradeoff between lower readability versus usability for some types of complex tasks involving a conjunction of data attributes? How can aesthetics and semantics be included in this framework? How can interactions – such as toggling an encoding on/off, or adjusting font size – enhance and make these techniques more usable? How can these techniques be evaluated with a wider range of datasets, at different scales and in different languages?

These questions imply that there is potential for much more research within this expanded design space. It continues to be an ongoing area of investigation for the authors.