Evaluation of Visualization by Critiques

Richard Brath
London South Bank University and Uncharted Software
brathr@lsbu.ac.uk

Ebad Banissi
London South Bank University
banisse@lsbu.ac.uk

ABSTRACT

In this position paper, we extend design critiques as a form of evaluation to visualization, specifically focusing on unique qualities of critiques that are different than other types of evaluation by inspection, such as heuristic evaluation, models, reviews or written criticism. Critiques can be used to address a broader scope and context of issues than other inspection techniques; and utilize bi-direction dialogue with multiple critics, including non-visualization critics.

CCS Concepts

• **General and reference ~ Evaluation** • **General and reference ~ Design** • **Human-centered computing ~ Visualization design and evaluation methods** • *Human-centered computing ~ Heuristic evaluations* • **Computing methodologies ~ Learning from critiques**

Keywords
Visualization Critiques; Heuristic Evaluation; Interdisciplinary Evaluation.

1. INTRODUCTION

Various researchers recognize that there is a significant design component in the creation of visualization systems (e.g. [1,2]). There are many different ways that a visualization can fail: for example, Munzner's nested model (e.g. [3]) identifies four levels with multiple evaluations per level. Yet, there are still many errors in visualization designs (e.g. [4]).

Instead, if it is a given that the creation of visualizations involves design, then the visualization community should consider evaluation approaches used in design. In particular, many types of design education use critiques as a form of evaluation used frequently throughout the design process. The contribution of this *position paper* is to expand on the use of critiques as a form of evaluation used in visualization design and show how it is different from other similar approaches such as heuristic evaluation, reviews or written criticism.

2. BACKGROUND

2.1 Evaluation and Design

Some types of evaluation techniques may be inadequate as they may not consider the many potential points of failure. For example, a visualization technique which focuses on pre-attentive perception of point marks, may achieve high performance on the time to perceive a target however, the encoding may not be easy to decode: a metaphoric or connotative encoding may perhaps be slower to perceive but faster to decode. Models such as Munzner's nested model, or Floridi and Chen's communications model [5] can be used to help identify different areas within a visualization that should be evaluated and adjusted. However, finer nuances are not necessarily captured by models such as a multitude of tradeoff decisions.

Evaluation by inspection is an approach to evaluating the broader visualization system and use. For example, heuristics have been compiled by different researchers (e.g. [4,6,7]). A heuristic evaluation focuses on judging a design to various established principles to assess the design's compliance to each heuristic and the approach has been used in visualization (e.g. [8]). As described by Nielsen [9], a heuristic evaluation is performed by having each evaluator inspect the interface alone – only after all evaluations have been completed are evaluators allowed to communicate and have their findings aggregated. The evaluator goes through the interface several times and inspects various user interface elements and compares them with a list of recognized usability principles. Heuristic evaluation does not provide a systematic way to generate fixes to the usability problems or a way to assess the probable quality of any redesigns [9]. Heuristic evaluation doesn't consider tradeoffs between different design choices (e.g. alternative encodings); assumptions associated with the heuristic that may not hold for the particular design; nor the possibility for conflicting heuristics. Tradeoffs can be more complex than simple functional and usability requirements: e.g. people are willing to trade aesthetics for functionality [10].

User-centered design is another approach for creating effective design. User-centered design is focused on user perceptions, behaviors, needs and experiences. The user-centered approach is focused on the problem space, but users "typically cannot directly articulate their analysis needs in a clear-cut way" [2]. Furthermore, as the authors have seen in past projects, some users may not be able to effectively provide feedback to visualization design ideas when expressed as wireframes or storyboards.

Models may discuss the need for iteration. For example, the nested model indicates that a better understanding of blocks at one level will feed back and forward into refining the blocks at over levels. While models can be very effective for framing the design process, models aren't inherently critical of the limitations of the models.

Following a user-centered approach, a model-based approach or a heuristic approach may lead to a workable solution. However, there may be better alternative solutions which are not necessarily found with these approaches. A user-centered approach is limited to user expertise – which may be low with regards to visualization. A model or heuristic approach is limited to the model constraints and knowledge-base of the designer or the heuristics. Even informal reviews with other visualization researchers may be constrained by models, guidelines and past findings. Past models, guidelines and

findings may not be universal to new types of problems, domains, technologies, user capabilities, assumptions, etc.

Design problems are characterized by having many potential solutions with tradeoffs between alternatives. The lead author has been involved in the design and development of many industrial visualizations. In one project, the key requirements provided by senior management included creating a visualization that showed a great amount of detail for thousands of named entities; that the representation be immediately comprehensible; that it should work in a tiny 200 x 300 pixel space. Further complications included incompleteness of the data and finding that users perceived the data as a hierarchy, but in fact there were multiple overlapping hierarchies, which were not necessarily strict hierarchies but rather directed acyclic graphs. The designer in this task has to confront various ambiguities and tradeoffs: Some data cannot be explicitly represented in such a tiny space and therefore some detail can only be accessed on interaction – but which content should be explicit and which should be interactive? How important is the representation of the hierarchy? Should the representation include a single hierarchy, multiple hierarchies or graph? How is it feasible to balance need for immediate comprehension (which implies simplicity) vs. need to show detail (which implies complexity)? Is it feasible to make named entities recognizable without explicit labels? And so on.

2.2 From Evaluation to Design Critique
The authors' position is that the discussion should be expanded beyond evaluation to a discussion to include design and idea generation. For example, Stuart Card in 2003 [11] says: "The rise in the dependence of HCI on usability labs is basically a regression… Design is where the action is. You will just never get great systems out of usability testing; you would never get to the GUI interface by usability testing on DOS." Or Don Norman in the same panel: "The design profession flourishes because they do things, they create. Usability languishes because good usability is invisible... Although we think we are indispensable, the world of business knows this to be false."

Donald Schön, in *Educating the Reflective Practitioner* [12], argues that most research universities are based on technical rationalism. Technical rationality holds that professional practitioners solve well-formed problems by applying theory from systematically derived scientific knowledge. However, real-world practice does not present well-formed problems, but messy indeterminate situations with a context often larger than the immediate requirements.

Instead, Schön argues for the constructionist view, wherein practitioners assemble *models rooted in perceptions, appreciations and beliefs which are continuously updated with new evidence from attention, sense-making, boundary-setting and so forth*. The designer's efforts (sketched and verbalized) provide the critic (i.e. practitioner) with evidence from which to infer the designer's difficulties and understanding forming a basis for the framing of questions, criticisms and suggestions. In effect, the critic is a coach.

Schön provides examples of this approach across many disciplines, including law, medicine, music, dance, art and architecture. Of medicine, he says: "There is an implicit recognition that *research based models of diagnosis and treatment cannot be made to work until the student acquires an art that falls outside the models*. The medical practicum is as much concerned with acquiring a quasi-autonomous art of clinical practice as with learning to apply research-based theory." [12]

Fred Brooks, lead developer of IBM's System/360 and winner of the Turing Award, supports Schön and says: "The weakness of much academic formal education is its reliance on lectures and readings, as opposed to critiqued practice… Only rarely do computer science curricula do that." [13]

Critique is not foreign to visualization: Kosara et al set out the basics of the design critique as applied to visualization in 2008 [14]. This position paper expands on the approach, to further differentiate critiques from other types of inspections such as heuristic evaluation, models, reviews and user feedback.

3. CRITIQUE

3.1 Critical Thinking, Criticism and Critique
Critique, criticism and critical thinking are closely related concepts. Critical thinking underlies both critique and criticism. Critical thinking is defined by Oxford dictionary as "the objective analysis and evaluation of an issue in order to form a judgement". Hughes et al say "Three types of skills – interpretation, verification and reasoning – constitute what are usually referred to as critical thinking skills."[15] Critical thinking will disassemble designs and models, question assumptions, reconsider evidence and hypothesize new models. Critical thinking is useful for open-ended questions with potential ambiguity and tradeoffs – questions with more than one right answer such as design problems. Critical thinking is self-guided, self-disciplined thinking. [16]

Unlike critical thinking, criticism and critique are explicitly public. Criticism originates in the 18th century during the Enlightenment when scholars and the bourgeoisie were struggling against absolutists in state and religion. It established a distinct public discourse based on rational judgement. Individuals gather for "equal interchange of reasonable discourse" in public forums such as clubs and coffee houses. [17] Criticism is "*open to debate, it attempts to convince, it invites contradiction*. It becomes part of the public exchange of opinion." [18].

In modern English usage, the word *criticism* tends to be associated with the publications of the professional critic, such as a movie critic or fashion critic: "The role of the serious critic is that of an educator. *By searching out the many examples of good design and appraising them constructively*, he may convince the manufacturer or printer of the merits of good design associated with his product... Such constructive criticism in the press would teach the public, not only to appreciate, but to demand good design in the products they buy." [19] Or, more to the point: "The critic has long been the arbiter of taste, determining for their readership what is considered good and what is bad." [20]. There are calls for increased criticism in visualization, such as the many examples on the blogs of Robert Kosara (eagereyes.org) or Kaiser Fung (junkcharts.typepad.com). Following in the model of the professional critic, this approach can lead to a better appreciation of good visualizations. However, discourse in written criticism occurs in slow motion: it is not a face to face dialogue in a coffee house, but rather unfolds with a slower batch-process carefully crafted opinion (and hopefully well-crafted responses).

As opposed to criticism, critiques (as used in education of design) are face to face interactions between designers and critics. The notion of critique can have subtle variations in meaning when applied to literature, philosophy or design. For the purposes of this paper, critique will be used in a design context and specifically refer to critique as used in the architectural design process.

3.2 Design Critique

The lead author has past experience in more than a hundred critiques through the completion of five years of undergraduate architectural education leading to a degree and two years of professional practice. This included experience at a variety of different architectural offices around the world and working with architectural students from other universities. In the last 20 years the lead author has worked in a visualization firm and used techniques borrowed from critiques to evaluate and advance design ideas.

Below is an outline of some of the unique qualities of design critiques with which the visualization educator or practitioner may be unfamiliar with. This is not an exhaustive review of the critique process: there exist more detailed analyses of architectural design critiques (e.g. [12]).

3.2.1 Sketches and design artifacts

Critiques are used frequently throughout the course of an architectural design project. In both architecture schools and professional practice, designers typically work in an open office (aka studio) continuously ideating, expressing and refining design ideas through sketching, physical models, virtual models, mockups, diagrams, illustrations, annotations and other representational embodiments of the design ideas.

Sketching, in particular, is a simple medium that can be utilized by *both* the designer and critic to reveal qualities and relations unimagined beforehand. These dynamic modifications function as quick exploratory experiments which are not restricted or slowed by real-world constraints [12].

Design critiques are applicable to user interfaces and visualizations, particularly where the design process similarly generates sketches, walkthrough, wireframes, Wizard of Oz prototypes and other visual artifacts, e.g. [21]. The notion of sketching has also been applied to visualization design, such as collaborative sketching [22], or as a method to generate design alternatives [23].

3.2.2 Broad scope

Given that design may have many tradeoff decisions with no single correct solution, a critique can be very wide ranging, including going beyond immediate functional requirements and may consider the elements of the broader social, historic, theoretic contexts. Following its Enlightenment origins, a critique is willing to question the underpinnings of the current theory and is willing to consider breaking rules.

A first year architecture student may be asked to design a simple building such as a house. A wide ranging critique will ask "why" about any aspect of the design. It may include aspects which are conceptual (why is it shaped the way it is, what governs the high level organization); spatial (there should be a bathroom near the bedrooms), regulatory (there are minimum sizes for some rooms, doorways, ceilings), structural (is a span too long and likely to need intermediate support), economic (a long span is more expensive than a short span), functional (the roof won't drain properly), visual (why are particular colors used), physiological (a door is too narrow), conventional (why a straight stair is the convention over a spiral), psychological (a low ceiling can induce a sense of compression), historical (how does the design relate to its historic context), social (how does it fit into the existing neighborhood), theoretical (is a motorhome a house, does a house need to include a kitchen), material (why are material choices made), and so on.

Similarly, visualization and visual analytics systems have many design inputs and design decisions to be made, including understanding user types, user capabilities, tasks, goals, workflows, data available, data types (literal, categorical, quantitative, free text, multimedia), data quality (nulls, certainty, provenance), data scale (kilobytes or terabytes), latency, analytics, models, encoding of visual variables (e.g. size, intensity, color, curvature, shape, texture, font), layouts (e.g. grids, recursive areas, springs), labels (axes, elements, titles), navigation (e.g. zoom/pan), probes (e.g. tooltips), collaboration, selection (e.g. click, tap, grasp), UI (e.g. buttons, sliders), animation (e.g. timeline, object constancy), speed of perception, accuracy of perception, ease of decoding, cognition and so on.

Furthermore, as noted by various authors (e.g. [24]), the theories of visualization are still evolving. The underlying science still has many gaps: we identify visual variables for visualization based on preattention research but rank visual variables based on accuracy of decoding. The list of visual variables varies per researcher [25]. There are many different tasks that visualization can be used for (analysis, monitoring, communication, ambience, etc), but visual variables aren't considered with respect to different uses. And so on. A critique is willing to consider design alternatives within these gaps, and explore beyond the current conventions whereas other approaches (heuristics, models, feedback) may be constrained by current best practices (e.g. [26]).

3.2.3 Unifying concepts and consistency

A design needs to define and follow some broad conceptual reasoning. The overall conceptual framework is important because smaller design decisions follow the larger rationale and make the design predictable and internally consistent. This internal consistency makes the design predictable and more efficient for the user: e.g. letters within a font have similar widths, x-heights and terminals, which facilitates reading [27]. Mies van der Rohe's Seagram's building in New York breaks with centuries of architectural tradition introducing uniform floor sizes and heights; unadorned structure and large glass windows allowing light deep into an office building coincident with concepts of modernity, technology, the rise of professional managers and democratization of the workplace (fig. 1).

Figure 1. Left: Seagram building, an exemplar of modern design (Mies van der Rohe, 1958) compared to Chrysler building, completed a generation earlier (William van Alen, 1930). (Public domain images from Wikipedia and Library of Congress)

Visualization systems may not have this internal consistency especially when cobbled together out of components. Perhaps such a system might meet functional requirements and heuristic checklists but do not have consistency of design. Different glyphs, encodings, sizes, styles, color, typography, layout, white space and interactions reduce the ability to take what is learned in one part of the application and use in another, e.g. [28].

User mental models may also be challenged by mixing representations and aggregations: a linked coordinated views visualization may include individual data elements explicitly represented as dots in a scatterplot, represented as lines within a parallel coordinate chart, and summarized into bars within a bar chart. A novice user of such a system may have greater cognitive load than the user a system with equivalent functionality wherein each data point is explicitly represented throughout, in scatterplots, stacked into bars and stacked into distributions.

3.2.4 Broad context and case studies
Suggestive alternatives discussed in a critique frequently cite other examples where a similar design problem may have had a unique, innovative approach to solving it. Historic examples and case histories, with many illustrative artifacts, are utilized as references for both the critic and the designer.

Regardless of the evaluation approach, there is value in the collection, curation and publication of exemplars [19]. This should not be limited to novel visualizations but also include examples of best practices in specific applications with suitable high-resolution screenshots and videos of workflows. These collections can inform future designers of relevant design criteria that may not be otherwise documented. Furthermore, visualization researchers and designers should also consider alternative sources of design influence including blogs, related conferences and workshops, and related visual design fields (e.g. information design, UX design).

Visualization also has a long historic context which in turn could influence design choices. This history goes beyond visualization conferences, Bertin's *Semiology of Graphics,* back to early pioneering examples such as Neurath's *Isotype,* or the various examples found in *Album de Statistique Graphique*, Gantt, Marey, Minard, Playfair and others; and information graphic techniques that have existed for more than a hundred years in adjacent domains, such as cartography, genealogical diagrams, organizational charts and financial charts (e.g. [29]). The lead author has overlooked relevant earlier work in practice at his peril: for example, within financial services, there are existing conventions for candlestick representations. Candlestick visualizations have existed for hundreds of years, but are not discussed within the visualization community.

Visualizations need to work within cultural preconceptions, metaphors, and codes of users, e.g. [30,31,32]. A critique can help identify possible unseen associations. One author's visualization had issues because it used gridlines which were not acceptable in one region of the world (i.e. gridlines are a distraction which add noise to the plot area of a chart reducing the ease of identifying patterns). Again, collections that include exemplars from around the world could have informed both designer and critic.

3.2.5 Public, bi-directional dialogue
Unlike a review of a paper, a critique is a dialogue between the designer(s) and the critic(s). The designer(s) may provide an overview of the design, an explanation of the rationale behind various design choices, defend various design decisions or suggest additional considerations. There is no anonymity for either side of the discussion.

Tom Hanrahan, Dean of Pratt School of Architecture says: "The great thing about architecture schools is that it takes place in a space where people discuss the work together, in a personal way (a one on one way) and in a very public way. Ultimately there's a public arena where the work is discussed, where students can present themselves, personally to other people, and show that they have a stake in the work and what they really think about the work." [33]

As noted earlier, criticism is open to debate, attempts to convince and invites contraction [18]. Rather than follow conventions, critique is willing to re-evaluate prior convictions and evidence in a different context. Critique can be effective because designers may become inadvertently locked into a particular design: e.g. they may be enamored with a particular design or feel that a design is the result of particular requirements and constraints. A critique provides an outside perspective which can challenge the designer, the constraints or the conventions. This helps reframe the problem and provides the potential for different design approaches.

In one visualization design project, the client had a multi-dimensional dataset. A key variable ranged across 6 orders of magnitude. With a bar chart, most values simply disappeared. Instead, the designers came up with two different design approaches. One design based on bars was dismissed as it didn't express small values well, even with variations using log-based axes and side-by-side bars. A second design was based on cubes, as volumes can express a higher dynamic range than lengths (e.g. bars) or areas (e.g. treemaps). The client dismissed this approach too. The designers were disappointed as they felt both designs presented novel ways to deal with the wide range of magnitudes. In a critique with a broader group of designers, the key question was "why did 6 orders of magnitude need to map to a size dimension (e.g. length, area, volume)?" The solution was redesigned using brightness intervals which was enthusiastically accepted by the users.

In another situation, two highly-published infovis experts were invited to critique a fully implemented infovis for a F500 client. One immediately dismissed the visualization as too complicated: too many visual attributes were in use simultaneously making it difficult to understand. The other immediately responded that the visualization was terrific. This critic disassembled the visualization and pointed out how multiple encodings aided the user to integrate many simultaneous data elements into a holistic view; and the metaphoric encoding reduced the cognitive effort. Much discussion was generated, in public, where each attendee could individually draw their own conclusions. Although the two views are opposing, both views have validity, and both views provide different directions as to how this particular visualization could evolve.

3.2.6 Many kinds of critics
A typical evaluation study might use novice users in a controlled experiment. Novice users, without expertise, only provide information on the task directly evaluated. Peers and experts, on the other hand, can provide feedback on any part of a system with which they are familiar, beyond the scope of the particular experiment: they may notice problems at different levels or different assumptions within the proposed design.

Criticism is not limited to a single expert. Multiple experts are involved throughout process and peers too. Experts can come from across related disciplines. From an architectural perspective, participants in critiques have included writers, historians, artists, theater designers, clients and engineers. Within the visualization

domain, experts could be sourced from the visualization community, the user community, or related fields (e.g. human-computer interaction, end-users' sales, service and IT departments, or sales or service department associated with end users, designers familiar with interface or information design, cartography, etc.) A broader range of participants can potentially help identify issues including less common issues.

3.2.7 Frequent

Design is a continuous, iterative process making frequent refinements throughout. Critiques are not once at beginning and once at end, but can occur frequently throughout a project.

Figure 2. Critiques in the design studio: with peers; with an expert at a desk critique; within a larger forum of peers and experts (Images from the documentary film Archiculture © 2016 Arbuckle Industries).

They may range from short ad hoc critiques by peers, to "desk crits" by from an expert, to a "final crit" by three or more experts in front of peers who also participate. In a one month design project, a student might receive four to eight desk crits from one to three different experts plus as many crits from peers as desired (figure 2). Frequent critiques are integral to an iterative learning process: an expert can criticize, from which the designer learns and creates new designs based on new knowledge, which in turn can generate new criticism. Frequent feedback provides greater opportunity to experiment with design alternatives, "fail-fast" and increase iteration cycles which can help reach a better design result.

3.3 Simple Design Critique Example in Viz

Consider the design of a small hypothetical visualization system for a financial fund manager looking at the performance of a few hundred investments. The fund manager is interested in the relative performance of investments per region (e.g. Europe, Asia, USA) and per industry (e.g. technology, industrials, consumer). The designer, after considering various alternatives, may select to use a 3D bar chart to represent this information. The designer may be aware of various authors who recommend against 3D (e.g. Tufte, Munzner), but may make a reasonable case for 3D by critically thinking through alternatives: e.g. a 2D grid-based layout with bars emphasizes rows or columns depending on bar orientation thus biasing perception; while a color-based heatmap does not provide the same dynamic range of values as a height-based encoding. A 3D bar chart can bypass both of these problems (figure 3).
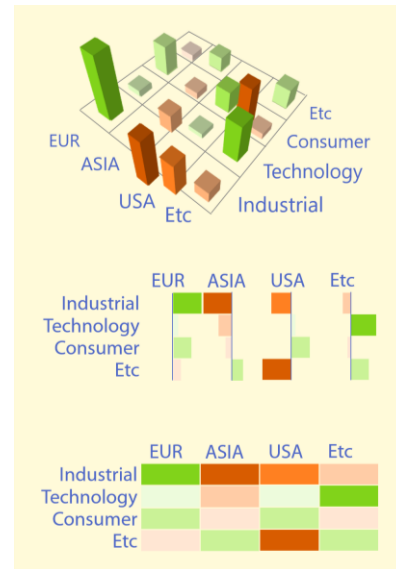

Figure 3. Potential design alternatives for cross-tabular quantitative data. The 3D variant provides greater range of values to be perceived than the heatmap and is not biased to promoting comparisons between rows or columns.

At an interim critique, an expert reviewer may or may not agree with the reasoning of the designer; and may suggest instead a matrix using bubbles at each intersection; question the potential use of thematic maps in the target domain; ask whether the task requires first assessing one split over the other (i.e. whether the task is better addressed with a hierarchical representation); whether a 3D projection should be isometric or perspective; whether performance over time is important and how time ranks to the other criteria; whether interaction is necessary or available in all the anticipated use cases and target devices; raise perceptual issues in 3D designs (e.g. occlusion) which in turn may be mitigated through various means including interactive navigation or spacing to aid perception of items; data characterization (is the data typically a normal distribution around zero, meaning very little occlusion); whether differentiation and/or identification of the individual elements is necessary; and so on. Note how the preceding critique hasn't included the myriad of other design factors such as color, labels, interaction, etc.

At a more formal critique with multiple reviewers, one reviewer may be skeptical and dismissive of 3D. Another reviewer, perhaps familiar with sources such as Munzner and Tufte, can defend the approach, addressing each objection in turn, for example: 1) full

occlusion may not occur if the initial viewpoint is such that items in the foreground do not completely cover items behind them or if the data tends to be have many values near zero height with just a few outliers; 2) the perspective distortion can be mitigated through the use of 2D projections such as isometric projections or the use of perspective cues (such as the base grid); 3) tilted text does not occur as the design choice uses horizontally aligned text; 4) height rather than depth is being used to encode key quantitative metric, and even adjusting for factors such as lower accuracy, the extra pixels in height gained by 3D might be able to provide higher accuracy than constrained 2D; 5) there exist examples of effective 3D visualization of abstract data (e.g. Munzner fig 6.8 [2], and 2014 IEEE VIS Workshop on 3DVis). Furthermore, the other critic may be invited to analyze the 3D vs 2D example provided in Munzner fig 6.4 [2] and consider the miscellaneous category: in the 2D charts one has to compare bar lengths for three bars in three different charts of almost equal value necessitating reading of the axis to answer the question, whereas it can be perceived immediately that the three bars are not equal length in the 3D view.

The important aspect of this sample critique is not the 2D vs 3D debate; rather it is the importance of debate, conviction and contradiction. Critics must be skeptics willing to question how far previous work can be generalized. Deep discussion is more important than strict adherence to rules, models and heuristics.

3.4 Do-it-yourself Visualization Critique

Given that visualization does not yet have a culture of critiques, it is possible to assemble a do-it-yourself (DIY) panel of critics. The lead author is currently pursuing a part-time, at-a-distance PhD in visualization. There are regular monthly meetings with the supervisor, but these represent only a single point of view: even heuristic evaluation techniques recommend a minimum of 3 reviewers [9].

The visualization designer can solicit advice, feedback and broad criticism from experts in vis and related fields as well as the target application domain. This can occur in formal venues, such as a doctoral colloquium, which unfortunately only occurs once during a doctoral research program. More flexibly, critiques can be solicited on an ad hoc informal basis, such as direct conversations (in person or via Skype); seminar presentations (or meetups) with significant Q&A portion with specific prepared questions for the audience; pre-arranged discussions at conferences with specific researchers; blogging and attempting to engage criticism via social media channels; indirect question and answer (e.g. Email); etc.

Over the course of the last 30 months, the lead author has solicited specific critical analysis of various portions of a PhD thesis (fig. 4) with 16 info vis experts, including three authors of visualization textbooks, five vis pioneers (authors of historic seminal research papers) and eight other vis researchers. Beyond vis, the author has solicited and received critiques from experts in HCI, financial services, typography and bloggers. Invited critics, unfamiliar with critiques may be very narrow in their responses: they can be coached with some positioning (e.g. "respond like a restaurant critic would respond to a restaurant"); and prepared questions can engage them to provide broader criticism, for example, to address how much the proposed solution addresses of the broader task, how the proposed solution might be utilized by a co-worker, etc.

Gaining feedback via critique can be an effective means to gain insight from adjacent communities with which the researcher is less familiar but relevant to the design research in question. For example, the author working on text visualization first read 10 books on typography, then spoke at a conference on typography where 20 different typographers were then engaged in one-on-one conversations reviewing in-progress design; as well as attended a graduate-level intensive typographic design course at a different university and engaged these experts as well.
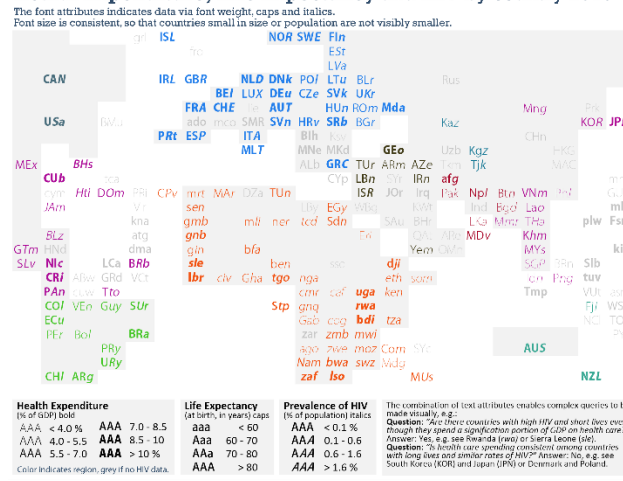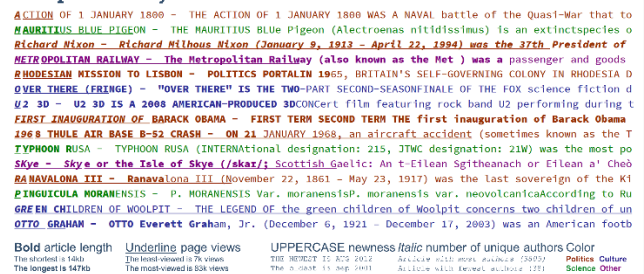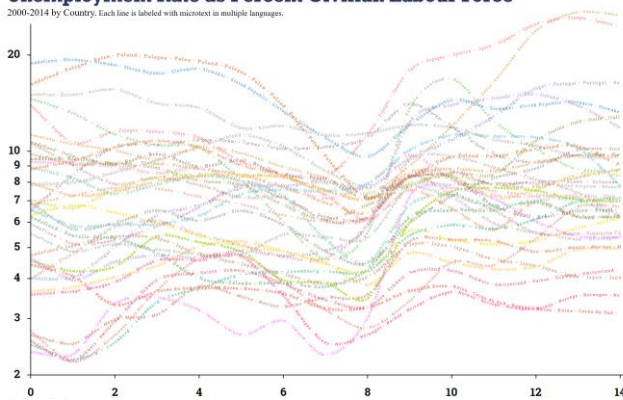

Figure 4. Snapshots of novel visualizations with unique typographic encodings, including thematic map with labels encoding multiple values; titles of articles representing quantities with formatting applied to a length of text; and a line chart wherein lines are replaced with text. What do typographers think? (Images © 2016 Richard Brath.)

By reaching out to a different domain, new issues were raised. For example, typographers immediately indicated legibility and readability of type are critical factors in typography, something unknown or not raised by any critiques with visualization researchers and absent in text visualization research (e.g. [34]). In general, reaching out to adjacent design communities may help improve the quality of visualization.

More broadly, by observing trends in criticism and critiques in other domains, we may gain insight into our own domain. For example, Ken Frampton (architectural historian at Columbia) critiques the romanticism of technology which can wrongly assume that the maximization of a technique is desirable: "Maximization is a problem altogether in architecture and other fields. Maximization of high tech surgery. Maximization of use of pharmaceuticals. Maximization of fertilizer to maximize agricultural production. Use the technology that's appropriate to the task."[35] While originating from architectural context, the critique is highly applicable to various visualization and analytic projects: some projects and papers utilize the latest or popular techniques, not because they are required, but because the technique is new, popular or just different. E.g. sometimes a table is better than a visualization [2], a summary is better than a tag cloud [36], a deterministic graph layout vs. a spring-based layout [37], and so on.

Experts are generally responsive (not everyone responds, some are slow to respond). Specific open-ended questions should be formulated in advance in order to start a conversation which can then be expanded in scope, assuming the critic is intrigued by the proposition. The conversation should start within an area of the critic's expertise as related to the researcher's current work. The researcher must first familiarize his/herself with the critic's area of expertise: don't expect a critic to spend time understanding your work, if you haven't taken the courtesy to first understand their work and how your work is related to their research.

4. DISCUSSION & RECOMMENDATIONS

There may be many objections to the use of critiques.

Science: Most visualization research and design in universities is associated with computer science and quantifiable scientific results are required. As Schön suggests, technical rationality has limitations. There are risks with evaluations that fit only within existing models and frameworks (e.g. bridge failure at Tacoma Narrows in 1940; or inadequate structure of Citicorp tower [38]).

It is the authors' position that if there is a design component to the creation of effective visualization, then vis will need to evolve beyond the confines of only researching what can be measured. Similar to HCI, visualization will be better served by a transdisciplinary perspective that honors both the rigor of what is measurable as well as the nuances and subtleties of that which is not measurable [39]. Critique is a valuable addition to evaluation "rather than forcing evaluations into experimental tasks that bear little resemblance to activities that most people routinely do." [40]

Lack of Experts: Unlike architecture, music or graphic design, there are far fewer visualization experts making it difficult to assemble a critique with three experts. This objection becomes weaker as more vis graduates are added to the field every year; and as graphic design programs add more information design and user interface education to their programs. Furthermore, looking only within visualization is severely limiting. Innovative solutions to similar problems may occur in human factors, graphic design, cartography, historic charts, typography, etc. Failing the ability to create a community for critique within an institution, there is the much larger community of experts, the potential to create a DIY critique, and the potential for experts at one university to critique at another university in a reciprocal arrangement.

Need for History and Case Studies: This is a gap. The bitmaps in most papers are tiny and difficult to decipher. The original visualization may no longer be operable e.g. [41]. There may be highly confidential data used. As a community, there is a need to identify better ways of documenting visualizations, their interactions, their use and how they handle various issues. Many interesting visualization system papers have been rejected because the author could not express what was the unique contribution of their system resulting in a rejection "there is nothing new here," in spite of the fact that it might be used within a unique user community, or that it is an exemplar of a particular technique or such.

With regards to documentation of systems, supplementary materials are nice, but it is difficult to browse and access these supplementary materials. They are not indexed or preserved with the same ease of access as papers. Blevis argues that visual content (in his case, photographic essays) can be on par with the textual content in research papers and of equal area rather than confined to tiny figures.[38]

Need for Design Rationale. Underlying concepts and design rationale are sometimes not captured in research papers. Visualization techniques are used with no justification, or perhaps a single sentence. There should be some documentation of the rationale and consideration if design assumptions match the assumptions of the referenced technique.

Need for Secrecy. One objection that is expressed to the DIY critique approach is that a researcher can't afford to divulge their work prior to publication for fear that the ideas may be copied and/or the risk that the researcher may not be first to publish the new idea. If the researcher has solicited criticism from many distinguished authors, the researcher will become known as having started the discussion. Secondly, design spaces are vast, permutations are near infinite: there may be many alternate solutions. Thirdly, secrecy may be related to patents or confidential information: non-disclosure agreements can be used in order to permit a conversation where secretive information is involved.

5. CONCLUSION

If objections can be overcome, then specific aspects of critiques should be used in visualization discourse. At the level of the individual designer, there should be more frequent engagement with a broader range of interdisciplinary critics, in a dialogue willing to challenge the status quo if needed. If local critics are unavailable, do-it-yourself critiques may be achievable.

Critics must be willing to be skeptics: of the design to be evaluated, of visualization models, studies and current best practices. At the level of the community, a broader collection of case studies should be assembled including examples of best practices in specific applications with suitable high-resolution screenshots and videos of workflows.

Future work could include consideration how critiques can be used together with other inspection methods and how the role of critiques can function beyond evaluation to more broadly support visualization idea generation. Critiques might also be used to more broadly foster interdisciplinary collaboration, feedback and evaluation: this can be used to expand the range of criteria that is considered in the evaluation including factors such as social, cultural, economic, sustainability, graphic design, and so on.

6. REFERENCES

[1] Paul, C. L., Rohrer, R., & Nebesh, B. 2015. A "Design First" Approach to Visualization Innovation. *IEEE Computer Graphics and Applications*, (1), 12-18.

[2] Munzner, T. 2014, *Visualization Analysis and Design*. CRC Press.

[3] Munzner, T., 2009. A nested model for visualization design and validation. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6), pp.921-928.

[4] Forsell, C. and Johansson, J., 2010, May. An heuristic set for evaluation in information visualization. In *Proceedings of the International Conference on Advanced Visual Interfaces* (pp. 199-206). ACM.

[5] Chen, M. and Floridi, L., 2013. An analysis of information visualisation. *Synthese*, 190(16), pp.3421-3438.

[6] Ware, C., 2012. *Information visualization: perception for design*. Elsevier.

[7] Zuk, T. et al., 2006, May. Heuristics for information visualization evaluation. In *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization* (pp. 1-6). ACM.

[8] Tory, M., and Möller, T. 2005. Evaluating visualizations: do expert reviews work? *Computer Graphics and Applications*, IEEE 25, no. 5 (2005): 8-11.

[9] Nielsen, J., 1994. *Usability engineering*. Elsevier.

[10] Pan, Y., Roedl, D., Blevis, E. and Thomas, J., 2015. Fashion thinking: Fashion practices and sustainable interaction design. *International Journal of Design*, 9(1).

[11] Shneiderman, B., Card, S., Norman, D.A., Tremaine, M. and Waldrop, M.M., 2002, April. CHI@ 20: fighting our way from marginality to power. In CHI'02 Extended Abstracts on Human Factors in Computing Systems (pp. 688-691). ACM

[12] Schön, D.A., 1987. *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. Jossey-Bass, San Franscisco, CA.

[13] Brooks, F. P.. 2010 *The Design of Design.* Addison-Wesley.

[14] Kosara, R., Drury, F., Holmquist, L.E. and Laidlaw, D.H., 2008. Visualization criticism. *IEEE Computer Graphics and Applications*, (3), pp.13-15.

[15] Hughes, W. Lavery, J. and Doran, K. 2015. *Critical Thinking: An Introduction of the Basic Skills. 7th ed.* Broadview Presss.

[16] Elder, L. *Defining Critical Thinking*. Website: https://www.criticalthinking.org/pages/defining-critical-thinking/766 (accessed June 11, 2016).

[17] Eagleton, T., 2005. *The function of criticism* (Vol. 6). Verso.

[18] Hohendahl, P.U., 1982. *The institution of criticism*. Cornell University Press.

[19] Havinden, A., 1952. Does Today's Criticsim Help Design? *Printing Review*, Winter 1952.

[20] Gerber, A. and Triggs T. 2007. in *Blueprint, The Magazine for Leading Architects and Designers*, Oct 2007, page 80

[21] Buxton, B., 2010. *Sketching user experiences*. Morgan Kaufmann.

[22] Craft, B. and Cairns, P., 2009, September. Sketching sketching: outlines of a collaborative design method. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology* (pp. 65-72). British Computer Society.

[23] Roberts, J.C., 2011. The Five Design-Sheet (FdS) approach for Sketching Information Visualization Designs. *Proc. Eurographics Education Papers*, pp.27-41.

[24] Goebel, R., 2014, January. A sketch of a theory of visualization. In *Information Visualization Theory and Applications (IVAPP), 2014 International Conference on* (pp. 218-221). IEEE.

[25] Brath, R. The design space of typeface. In *Proceedings of the 2014 IEEE Symposium on Information Visualization (VisWeek 2014). 2014.* http://bit.ly/29TUa0F

[26] Santos, B.S., Ferreira, B.Q. and Dias, P., 2016. Using Heuristic Evaluation to Foster Visualization Analysis and Design Skills. *Computer Graphics and Applications*, IEEE, 36(1), pp.86-90.

[27] Gauthier, I., Wong, A., Hayward, W., and Cheung, O. Font tuning associated with expertise in letter perception. 2006. *Perception* 35, no. 4: pp. 541-559.

[28] Jonker, D. et al. 2013. Aperture: An Open Web 2.0 Visualization Framework. In System Sciences (HICSS), 2013 46th Hawaii International Conference on, IEEE, pp. 1485-1494.

[29] Friendly, M., and Denis, D. *Milestones in the History of Thematic Cartography, Statistical Graphics and Data Visualization: An Illustrated Chronology of Innovations.* http://datavis.ca/milestones/ (accessed June 12, 2016).

[30] Norman, D. 2002. *The Design of Everyday Things*. Basic Books.

[31] Byrne, L., Angus, D. and Wiles, J., 2016. Acquired Codes of Meaning in Data Visualization and Infographics: Beyond Perceptual Primitives. *Visualization and Computer Graphics, IEEE Transactions on*, 22(1), pp.509-518.

[32] Jones, J., 2015. Information Graphics and Intuition Heuristics as a Techne for Visualization. *Journal of Business and Technical Communication*.

[33] Hanrahan, T. 2014. in *Archiculture* (documentary film). Arbuckle Industries, Dec 17, 2014. 04:38. https://www.youtube.com/watch?v=62r3UPrOS9k (accessed June 8, 2016)

[34] Strobelt, H. et al. 2016. Guidelines for Effective Usage of Text Highlighting Techniques. *Visualization and Computer Graphics, IEEE Transactions on*, 22(1), pp.489-498.

[35] Frampton, K. 2014. *Ken Frampton Archiculture Interview*. Arbuckle Industries, Dec 17, 2014. 04:00 https://www.youtube.com/watch?v=NzZ1VsrUyXU

[36] Nielsen, J., 2009. *Tag Cloud Examples*, Nielsen Norman Group, last modified March 24, 2009. https://www.nngroup.com/articles/tag-cloud-examples/

[37] Kieffer, S. Dwyer, T. Marriott, K. and Wybrow, M. Hola: Human-like orthogonal network layout. *IEEE Transactions on Visualization and Computer Graphics,* 22(1), pp349-358.

[38] Morgenstern, J. 1995. The Fifty-Nine-Story Crisis, *The New Yorker.* May 29, 1995, pp 45-50.

[39] Blevis, E., 2016, June. Being Photo-Visual in HCI and Design. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems* (pp. 983-995). ACM.

[40] Dillon, A., Reading from Paper versus Screens: A Critical Review of the Empirical Literature. *Ergonomics,* 35(10). pp. 1297-1326.

[41] Kosara, R. *The Bits Are Rotting in the State of Data Journalism*. July 13, 2016. http://bit.ly/29QArh1