Beyond Usability and Performance: A Review of User Experience-focused Evaluations in Visualization

Bahador Saket
Georgia Institute of Technology
saket@gatech.edu

Alex Endert
Georgia Institute of Technology
endert@gatech.edu

John Stasko
Georgia Institute of Technology
stasko@cc.gatech.edu

ABSTRACT

Traditionally, studies of data visualization techniques and systems have evaluated visualizations with respect to *usability goals* such as effectiveness and efficiency. These studies assess performance-related metrics such as time and correctness of participants completing analytic tasks. Alternatively, several studies in InfoVis recently have evaluated visualizations by investigating *user experience goals* such as memorability, engagement, enjoyment and fun. These studies employ somewhat different evaluation methodologies to assess these other goals. The growing number of these studies, their alternative methodologies, and disagreements concerning their importance have motivated us to more carefully examine them. In this article, we review this growing collection of visualization evaluations that examine user experience goals and we discuss multiple issues regarding the studies including questions about their motivation and utility. Our aim is to provide a resource for future work that plans to evaluate visualizations using these goals.

Keywords

Data visualization, user experience goals, usability goals, memorability, enjoyment, engagement.

1. INTRODUCTION

Evaluating what has been built is at the heart of a design process [42]. Evaluating a design allows one to ensure that it is appropriate and meets user requirements. For many years the notion of a "successful" design has centered on a system meeting all of its *usability goals* [41, 42]. Usability refers to the effectiveness, efficiency, safety, utility, and learnability of a design [17]. These principles have been at the core of evaluation research within the field of human-computer interaction for decades.

Today, however, not only designers but also cognitive scientists acknowledge that other important design goals must be considered. These new goals go beyond traditional usability-driven objectives and address issues related to a person's experience when using a system. In a world full of choices where the short attention of a person becomes a prime resource, it is essential that designers create not just usable products, but also memorable, fun, enjoyable, and engaging experiences [41]. Multiple HCI researchers have emphasized the importance of emotion, enjoyment and fun, memorability, and engagement in their work [5, 19, 37, 55, 56]. As Rogers et al. advocate, designs should meet both *usability goals* and *user experience goals* [54]. User experience goals complement traditional usability goals to provide more holistic evaluations. In his recent book *Emotional Design*, Don Norman states [41]:

> *" we scientists now understand how important emotion is to everyday life, how valuable. Sure, utility and usability are important, but without fun and pleasure, joy and excitement, and yes, anxiety and anger, fear and rage, our lives would be incomplete."*

Like that done in HCI, the field of data visualization also has embraced the importance of usability. However, the primary design goal for visualization is to effectively communicate a thorough understanding of the data it represents. This utility of a visualization does include usability goals, but ultimately revolves around the visualization's ability to help people better understand data. We will use the term "performance" to represent this ultimate purpose of a visualization. The vast majority of visualization user studies and evaluations thus have focused upon performance-related characteristics, and they have assessed metrics such as task completion-time and accuracy, which can be objectively measured.

The BELIV Workshop was created to explore new evaluation methods that go beyond these well-established metrics such as time and errors. While it has largely succeeded in doing so, we would argue that the vast majority of research presented at the workshop has focused on performance-related metrics such as "insight" or has proposed new evaluation methodologies such as more observational, qualitative case studies. Those metrics and methods still relate to a visualization's ability to help people understand data better and answer questions about that data. Evaluation studies of this type still focus on assessing performance goals.

As discussed above, the HCI community has now embraced the importance of design goals beyond the usability of a system. Recently, a similar effort to encompass and evaluate design goals beyond a system's prime purpose has emerged in the field of data visualization as well. In visualization, multiple projects have evaluated aspects of a visualization beyond its utility and performance-related characteristics. Several recent papers study enjoyment, fun, memorability, and engagement of visualizations. Brehmer and Munzner [12] recognize enjoyment as one of the three reasons that one might use visualizations. Haroz et al. [24] argue that a primary goal of some visualizations (particularly news articles) is to engage people and make them pause and look. Several studies also emphasize the importance of memorability in visualizations [9, 8, 52]. These types of studies that move beyond performance-related objectives are not universally accepted, however. Some people question whether they

BELIV '16 Baltimore, Maryland USA

DOI: 10.475/123_4

deserve a place in the visualization research agenda and believe they only distract from the most important cognitive measures [21].

Historically, the general interest in and adoption of usability goals can be explained at least in small part by how well-understood and well-defined these metrics are to measure (e.g., performance time and accuracy). This is especially true in contrast to user experience goals that appear to be more difficult to define, much less to measure [50]. Despite this challenge, a number of studies recently have begun to examine and evaluate visualizations with respect to user experience goals. However, the methods these studies applied vary significantly, employing qualitative methods, quantitative methods, or some combination of both. Unfortunately, no existing work summarizes and classifies the methodologies used by these studies to assess user experience goals.

In this paper, we review a collection of academic publications and online articles that evaluate visualizations based on user experience goals. Each of them contribute methods, metrics, and/or comment on the application of user experience goals in data visualization. More specifically, we concentrate on memorability and recall, engagement, and enjoyment and fun (defined in Table 1). In reviewing the existing literature in this area, we specifically explain the methodology used by each study to measure these metrics.

2. USER EXPERIENCE GOALS

Undoubtedly, the vast majority of people would agree that a visualization's ability to convey understanding of the data it represents is absolutely paramount. However, researchers have now begun to explore other applications and benefits of visualizations as well. The designers of visualizations often have different goals for their creations. Accordingly, how one evaluates a visualization should depend on those corresponding goals.

In this section we review user studies that have focused on goals and metrics other than typical performance measures related to knowledge about the underlying data. We examine three primary user experience goals: memorability and recall, engagement, and enjoyment and fun.

Our methodology for collecting articles reviewed in this work began by searching for academic publications that evaluated visualizations via these user experience goals. We looked in well-known visualization venues such as the IEEE InfoVis, PacificVis, EuroGraphics EuroVis, and ACM CHI Conferences, as well as the ACM BELIV Workshop. Additionally, we searched in the *Information Visualization* and *IEEE TVCG* journals. Within each relevant article we found, we examined the references to discover other potentially relevant articles in these and other venues.

While reviewing these articles we performed two phases of coding. In the first phase, we categorized each study based on the user experience goals (memorability, enjoyment, and engagement). In the second phase, we coded memorability studies based on what types of memory (e.g., short-term, working, and long-term) they evaluated. In both phases, we used the terms reported in each article to categorize the evaluation and report on the type(s) of memory being examined. That is, we categorize each memorability study by the type of memory it examines, according to the paper's authors. Note that in some cases different authors apply these terms for types of memory in different ways.

2.1 Memorability and Recall

Memorability has been defined as a capability of maintaining and retrieving information [13]. Multiple studies have examined the memorability of visualizations, a person's ability to remember and recall information about the visualization. These studies all have employed an implicit assumption that being able to remember a

Table 1: Definitions of the three user experience goals discussed in this paper.

GOALS	DEFINITION
Memorability	Memorability is a capability of maintaining and retrieving information [13].
Engagement	Emotional, cognitive and behavioural connection that exists, at any point in time and possibly over time, between a user and a resource [2].
Enjoyment	Feeling that causes a person to experience pleasure [16]. Pleasure is recognized with occurrent happiness and excitement, which can be explained in terms of belief, desire, and thought.

visualization is good. Presumably, this is because the visualization has made a lasting impact and facilitates a person revisiting or discussing the relevant data at a later time. Different studies have focused on the memorability of different aspects of the visualization, however, such as its message, the underlying data, or its visual imagery. Table 2 shows a summary of the methods these studies used to assess memorability. We also discuss each in more detail below.

2.1.1 *Memorability of Embellished Visualizations*

The work by Bateman et al. [5] was one of the first studies that aimed to measure memorability of visualizations. The goal of the study was to evaluate the impact of embellishments on visualization memorability and comprehension. The researchers conducted a multi-phase controlled experiment to test comprehension and recall of visualizations using embellished and plain versions. In the first phase of their experiment, participants were shown a series of visualizations and asked to perform four tasks (e.g., subject, values, trend, value message). In the second phase (recall phase), the researchers measured both short-term (5 minutes after the first phase) and long-term (about a week after the first phase) recall of the visualizations. Participants first were asked to remember as many visualizations as possible. They then were asked to describe the charts as completely as possible. The researchers found that the participants were better able to recall the embellished visualizations, and participants' accuracy in describing the embellished visualizations was no worse than for plain ones.

Li et al. [34] reported a replication of the Bateman et al. study, limiting their selection to charts of data sets with 10 or more observations. The researchers found that the presence of a time limit decreased the accuracy of participants in describing and recalling visualizations, while the type of visualization significantly affected short-term recall.

Borgo et al. [7] also conducted an experiment to examine the effects of visual embellishments in visualization processes in relation to several fundamental aspects of perception and cognition including both working memory and long-term memorability.The researchers showed 18 visualizations (6 horizontal bar charts, 6 vertical bar charts and 6 bubble charts) to participants. Some of the visualizations were embellished and some were not. First participants were shown each visualization for 9 seconds. Participants next were shown a gray masking screen for 5 seconds and then were asked a question regarding the visualization that just was shown. Study results indicate that when embellishments are grouped with a numerical representation, they have the most beneficial influence on working memory tasks. The researchers also investigated the effect of embellished charts on long-term memory. Participants were each shown a visualization for 9 seconds, similar as done in

Table 2: Different methods used to measure memorability in visualizations. Each row of the table describes a study that has measured memorability. Columns indicate WHAT each paper investigates in the experiment, HOW it was evaluated, WHEN it was evaluated during the experiment, HOW MANY subjects have participated in the study, and WHICH VISUALIZATIONS were used in the study.

REFERENCES	WHAT?	HOW?	WHEN?	HOW MANY?	WHICH VISUALIZATIONS?
Bateman et al. [5]	Immediate memory Long-term memory	**Immediate Memory:** Asked participants to recall as many visualizations as possible. **Long-term Memory:** Asked participants to recall as many visualizations as possible.	**Immediate Memory:** Five minutes after performing tasks using those visualizations. **Long-term Memory:** Two to three weeks after performing tasks using those visualizations.	20 participants	Bar chart, line chart, and pie chart
Li et al. [34]	Short-term memory Long-term memory	**Short-term Memory:** Asked participants to recall as many visualizations as possible. **Long-term Memory:** Asked participants to recall as many visualizations as possible.	**Short-term Memory:** Five minutes after performing tasks using those visualizations. **Long-term Memory:** Four days after performing tasks using those visualizations.	15 participants	Bar chart, line chart, and pie chart
Saket et al. [52]	Short-term memory Long-term memory	**Short-term Memory:** Asked participants to recall data shown in the visualizations. **Long-term Memory:** Asked participants to recall data shown in the visualizations.	**Short-term Memory:** Right after performing the tasks using those visualizations. **Long-term Memory:** Four days after performing the tasks using those visualizations.	40 participants	Node-Link and Node-Link-Group (map) visualizations
Borgo et al. [7]	Sensory memory Working memory	**Sensory Memory:** Asked participants to answer questions about visualizations. **Working Memory:** Asked participants to answer questions about visualizations.	**Sensory Memory:** Each visualization was shown for 9 seconds (5 seconds break between each two visualizations). **Working Memory:** Each visualization was shown for 9 seconds (30 seconds break between each two visualizations).	35 participants	Bar chart and Bubble chart, and line chart
Haroz et al. [24]	Working memory	**Working Memory:** Asked participants to recall data shown in the visualizations.	**Working Memory:** Each visualization was shown to participants for 1.5 seconds.	22 participants	Different variation of bar charts
Borkin et al. [9]	Not specified	Asked participants to press a key if they saw a visualization for a second time.	Visualizations were shown for a few seconds with 1.5 seconds gap.	261 participants	Almost all types of well-known visualizations
Borkin et al. [8]	Not specified	**Recognition Phase:** Asked participants to press a key if they saw a visualization for a second time. **Recall Phase:** Asked participants to describe visualizations as detail as possible.	**Recognition Phase:** After showing each visualization for 10 seconds. **Recall Phase:** After showing all visualizations (each visualization was shown for 10 seconds).	33 participants	Almost all types of well-known visualizations
Marriott et al. [38]	Short-term memory	**Short-term Memory:** Asked participants to redraw the visualizations.	**Short-term Memory:** After seeing and performing tasks using visualizations.	25 participants	Graph visualizations

the working memory section. Next, a gray masking screen appeared for 30 seconds. The researchers found that visual embellishment enhances information retention in terms of both accuracy and time required for short-term memory recall. Since they concentrated on "visual perception and cognitive speed-focused tasks" that leverage cognitive abilities, they used analytical tasks that forced attention to switch from one task to another.

Borkin et al. [9] conducted an online study with Amazon's Mechanical Turk that used thousands of real world visualizations. Their goal was to reduce the biases caused by the limited number of participants and target visualizations. Each target visualization was coded by the research team for different characteristics (visualization type, number of colors used, data-ink ratio, etc.). Their experiment was set up as a game on Amazon Mechanical Turk, where participants were presented with series of visualizations (each visualization was shown to participants for 1 second, with a 1.4 seconds gap between consecutive visualizations). Participants had to press a key if they saw a visualization for the second time in the sequence. The study found that attributes such as color and the inclusion of human-recognizable objects improve memorability.

Haroz et al. [24] tested the effect of pictographical representations on working memory (memory for briefly glanced information), as well as memory under load (memory is more crowded). In order to test the impact of pictographical representations on working memory, they ran a study where participants were shown visualizations with different values for 1.5 seconds. After 1.5 seconds, participants were shown the visualization again but this time values in the visualization were invisible. Participants were asked to recall each of the values in the visualization they just saw. In the second phase of the study, the researchers tested how pictographs impact memory when memory is more crowded. The procedure for this experiment was similar to the first phase (working memory) but used a "1-back design". Participants were asked to recall visualizations, but this time they were always tested on the visualization before the one that they just saw, introducing the need to store two charts at all times. The researchers found that ISOTYPE visualizations were recalled more accurately than those having plain charts.

More recently, Borkin et al. [8] conducted a more comprehensive study to move beyond memorability and investigate how visualizations are recognized and recalled. Their experiment contained three phases: *encoding*, *recognition*, and *recall*. In the encoding phase, participants were shown different visualizations (each for 10 seconds) while being eye tracked. The goal of this phase was to "allow participants more time to explore visualizations and to facilitate the collection of eye movements." In the recognition phase, participants were shown exactly the same visualizations as in the first phase of the experiment, as well as different unseen visualizations. Each visualization appeared for 2 seconds. Participants had to indicate which visualizations they remembered from the first phase of the experiment. Finally, in the recall phase of the experiment, participants were shown the visualizations they successfully recognized in the prior phase, and they were asked to describe the visualization in as much detail as possible. From this study, four general design guidelines emerge. First, visualizations that are memorable at a glance also have memorable content. Second, titles and text are important elements for recalling a visualization's message. Third, pictorial images do not decrease the memorability or understanding of a visualization. Fourth, redundancy increases recall and understanding of visualizations.

2.1.2 Memorability of Network Visualizations

Network diagrams, in particular, have served as the focus of multiple memorability studies [29, 38, 52]. As a part of an experiment comparing the effectiveness of four different visualization techniques, Jianu et al. [29] asked participants to complete several different tasks including one memorability related task. Their results suggest that Map-like network visualizations are more memorable than three other techniques. In a study aiming to investigate cognitive impact of different layout features (e.g, symmetry and alignment) on the memorability of graphs, participants were shown different graphs and were asked to study, remember and redraw them [38]. Their findings show that visual features such as symmetry, collinearity and orthogonality of network visualizations have an impact on memorability.

In another memorability focused study, Saket et al. [52] assessed the memorability of displayed network data, minutes and days after interaction with two different types of visualizations (node-link diagrams and map-based visualizations). The researchers argued that, unlike previous work, their study sought to measure the memorability of data shown in a visualization and not just the visualization appearance itself. They ran a three-phase experiment. In the first phase of the study, participants were asked to examine the visualization without a preset time limit and without performing specific tasks. In the second phase, participants were asked to perform different tasks using the visualization as fast and as accurately as possible. Finally, after a predetermined period of time depending on whether the subject was in the immediate or long-term treatment condition, participants were asked to recall the underlying data presented using visualizations. Study results found that participants recalled data in map-like visualizations more accurately, but not any faster, both in terms of remembering graphs and actual content.

In general, results of all these studies examining memorability have shown that embellishing visualizations appears to enhance their memorability. Almost all these studies additionally found that visualization embellishment did not come at the cost of performance. Since the majority of these studies evaluated the memorability of visual encodings used in visualizations and not the underlying data, whether embellishment enhances the memorability of the underlying data remains an open problem. In spite of these findings, the importance of the memorability of a visualization is not a universally agreed upon point. In fact, it has been the subject of heated debate recently. We will return to a discussion of such issues later in this article.

2.2 Engagement

User engagement has been defined as the *"emotional, cognitive and behavioural connection that exists, at any point in time and possibly over time, between a user and a resource* [2]". This broad definition of user engagement emphasizes its holistic character and different methods that can be applied to measure it. Previous work in visualizations refers to user engagement as users' interest in putting effort to investigate and explore visualizations and gain more insights [24, 10]. This definition of engagement supports the argument that an engaging activity attracts user attention to the activity, excluding other things and people [43]. The amount of time a person spends on an activity has been shown to be an effective indicator of the level of his/her cognitive involvement in that activity [44]. That is, the more engaged someone is in an activity, the more likely they underestimate the passage of time [3]. Read et al. [49] also indicated that people will remember engaging experiences and want to repeat them. This aspect of user engagement refers to the possibility of remembering an experience and the desire to repeat it.

Engagement factors are known as a medium for attracting users' interest in designs [41]. We speculate that this also applies to visualizations. Heer et al. [26] advocated for engaging new audiences for visualizations. Mahyar et al. [35] also suggested moving toward engaging visualizations.

Today, many visualizations are deployed on the Web where metrics such as visits and dwell time are considered important. Such metrics often drive commercial rewards and monetary gains. Fundamentally, a more engaging visualization that is viewed more often will be perceived as being more "successful" and impactful.

Table 3 summarizes different studies that have examined engagement as a primary metric, and we discuss each of these studies further below.

Haroz et al. [24] conducted a study to measure the engagement of different pictographical representations. More specifically, they wanted to know *"Will an ISOTYPE visualization be better at capturing attention than a simple bar chart?"* In the engagement phase of their experiment, participants were shown a 3x3 grid of items (3 bar charts, 3 stacked pictograph charts, and 3 pieces of text). Each item contained a short title above a small, blurred thumbnail. The thumbnail was either a visualization or a set of sentences about the topic. Participants were given two minutes and asked to look through the thumbnails. They could click whichever item they were interested in to view the information at full screen resolution. Clicking again would return them to the main grid. The researchers measured engagement of visualization by computing the total time participants spent working with each type of visualization. Study results showed that participants were more engaged with ISOTYPE-style visualizations.

Previous work also showed how product reaction card methods can be used to evaluate user experience with different visualization types [6]. Reaction cards provide "a way for users to tell the story of their experience, choosing the words that have meaning to them as triggers to express their feelings - negative or positive - about their experience" [4]. Suggesting the use of product reaction card method to evaluate user experience, Merčun evaluated four different visualization techniques (Indented list, Radial tree, Circlepack, Sunburst) to demonstrate how the results of this method could be analysed and used for comparing different designs [39]. Merčun first recruited 120 participants and asked each of them to work with three randomly counterbalanced visualization techniques. After completing 10 tasks

Table 3: Different methods used to measure engagement in visualizations. Each row of the table describes a study that has measured engagement. HOW **engagement was evaluated,** WHEN **it was evaluated during the experiment,** HOW MANY **subjects have participated in the study, and** WHICH VISUALIZATIONS **were used in the study.**

REFERENCES	HOW?	WHEN?	HOW MANY?	WHICH VISUALIZATIONS?
Haroz et al. [24]	Measured total time spent working with charts.	While participants were interacting with the charts.	10 participants	Different variation of bar charts
Boy et al. [10]	Measured total time spent working with charts. In addition, recorded interaction logs (e.g., number of clicks) with charts.	While participants were interacting with the charts.	Not included	Variety of interactive visualizations
Merčun [4]	Product reaction card method.	While participants were interacting with the charts.	120 participants	Indented list, radial tree, Circlepack, and Sunburst
Saket et al. [51]	Total time spent looking at a chart.	While working with the charts.	32 participants	Node-Link and Node-Link-Group (map) visualizations

using each visualization, participants were asked to choose from a set of cards (adjectives) those that best reflect their experience with the visualization. Finally, participants were asked to explain more about the visualization they used by applying the selected adjectives. The study found that participants were more engaged with sunburst and circlepack visualizations than other ones. In addition, the study shows that product reaction card method can be used to evaluate aspects such as engagement, ease of use, appeal, efficiency, and usefulness of visualizations.

Boy et al. [10] investigated the effect of using initial narrative visualization techniques and storytelling on user engagement with exploratory information visualizations. They ran three field study experiments. For each experiment, they created two different exploratory visualization webpages. One webpage contained *"an introductory narrative component, which told a short 'story' about the topic and context of the data provided initial insights and unanswered questions, and introduced the different visual encodings"*; and another webpage that did not. The researchers sought to understand whether augmenting such a visualization with an introductory 'story' can help engage users in exploration. In order to measure user engagement, they examined different user interaction logs such as the amount of time participants spent to explore, number of meaningful interactions, etc. Their results indicated that combining exploratory visualizations with introductory 'stories' does not enhance user-engagement.

Saket et al. [51] measured people's engagement with two different visualizations of the same relational data–node-link and map-based visualizations–by measuring the total amount of time participants spent looking at visualizations. Their main insight for this way of measurement is that *"if users are surreptitiously given the option to experience either of two visualizations, they will spend longer with the one they enjoy better."* In the first phase of their study, they invited each participant to a room where posters of two different types of visualizations were placed on two walls. The interviewer then told each participant: *"I need to bring some equipment before running the experiment. Please stay in the room for few minutes. I will be back soon and we'll start the experiment."* The interviewer then left the room and came back after a few minutes. During this time a camera was recording participants' interactions with the posters. The study findings show that the participants spent on average more time looking at the map visualizations but the difference was not significant.

Mahyar et al. [35] emphasize the lack of a taxonomy for measuring engagement in visualization, one that evaluates engagement based on the level of a user's understanding of a visualization. The researchers proposed a five level taxonomy for engagement. Each level of the taxonomy is correlated to the level of cognitive task user required to perform.

Expose: knowing how to read a data point
Involve: interacting with the visualization and manipulating the data
Analyze: analyzing the data, finding trends, and outliers
Synthesize: being able to form and evaluate a hypothesis
Decide: being able to draw final decisions based on evaluations of different hypotheses

These few prior studies of engagement used metrics such as a count of user interactions with a visualization and time spent viewing and/or working with a visualization. The studies were different enough so that no one common take-away message emerged from them. Although their methodologies can begin to help people measure engagement broadly, the metrics they assessed fall short of evaluating deeper levels of engagement. Additionally, it remains an open question of how to extract factors that make one visualization more engaging than another.

2.3 Enjoyment and Fun

Davis defines enjoyment as a feeling that causes a person to experience pleasure [16]. Pleasure is recognized with occurrent happiness and excitement, which can be explained in terms of belief, desire, and thought. Furthermore, enjoyable and fun experiences make people smile [41, 55]. Enjoyment has been carefully studied in psychology and HCI. For instance, Brandtzæg et al. [11] suggested that when designing for enjoyment, one should consider demands, allow a high degree of decision latitude, and provide socially rewarding activity. One of the best known models for understanding and measuring enjoyment in psychology is the flow model of Csikszentmihalyi [15]. In a series of experiments in different countries, people were asked to describe when and how they achieved the highest level of enjoyment when performing some activity. Csikszentmihalyi notes,

> *"Regardless of culture, social class, gender or age, the respondents described enjoyment in very much the same way. What they did to experience enjoyment varied dramatically – the elderly Koreans liked to meditate, the teenage Japanese liked to swarm around in motorcycle gangs – but they described how it felt when they enjoyed themselves in almost identical terms"*.

He then suggests several factors (challenge, focus, clarity, feedback, control and immersion) that encompass the experience of enjoyment. His flow model has been applied by researchers in other fields to create new models [23, 56, 57] and to assess enjoyment [45,

Table 4: Different methods used to measure enjoyment in visualizations. Each row of the table describes a study that has measured enjoyment. HOW **enjoyment was evaluated,** WHEN **it was evaluated during the experiment,** HOW MANY **subjects have participated in the study, and** WHICH VISUALIZATIONS **were used in the study.**

REFERENCES	HOW?	WHEN?	HOW MANY?	WHICH VISUALIZATIONS?
Bateman et al. [5]	Self Report (asked participants to rate enjoyability of each type of chart)	After performing all tasks using each type of charts.	20 participants	Bar chart, line chart, and pie chart
Li et al. [34]	Self Report (asked participants to rate enjoyability of each type of chart)	After performing all tasks using each type of charts.	15 participants	Bar chart, line chart, and pie chart
Saket et al. [51]	Applied Flow Model in visualizations [50] and asked open-ended questions.	After performing all tasks using each type of charts.	32 participants	Node-Link and Node-Link-Group (map) visualizations

46, 48]. Malone explored the topic of enjoyable and fun interfaces in the early studies of games [36]. He summarized the design heuristics for enjoyable and fun interfaces with these criteria: challenge, curiosity, and fantasy. Later, Shneiderman commented on enjoyable and fun experiences. He stated, *"They [enjoyable and fun experiences] are a break from the ordinary and bring satisfying feelings of pleasure for body and mind."* [55]

While enjoyment and fun have been discussed and studied extensively in psychology [11, 15, 48] and HCI [23, 37, 55, 56], the enjoyment and fun aspects of visualizations have not been explored significantly, even though enjoyment is often given as a reason to consume visualizations [12]. Moreover, previous work indicated that positive mental states appear linked to better problem-solving performance in general [22] and in information visualization in particular [25, 28]. In other words, if people are happier, then they will be more effective performing tasks using visualizations.

Fundamentally, people gravitate toward activities that are fun and enjoyable. As discussed above regarding engagement, metrics such as page visits and time on task are often viewed as measures of success, especially on the web. Online visualizations such as the Baby Name Wizard [1] have drawn tremendous interest and attention. The creators of the system speculated that factors such as enjoyment and engagement helped foster that interest [58].

Although enjoyment and engagement are actually two different concepts [59], some studies in visualization have used the terms interchangeably. We view them in this context as related, but not quite the same. Enjoyable activities typically are strongly engaging too, but it is clearly possible to be deeply engaged in an activity without the activity necessarily being fun and enjoyable. Table 4 reviews studies examining enoyment issues, and we describe each of the studies in further detail below.

Earlier, we described a study by Bateman et al. [5] to evaluate the comprehension and recall of charts using an embellished version and a plain version. As a part of this study, the researchers also asked participants to rate the enjoyability of each type of chart. Their results suggest that embellished charts are more enjoyable than plain ones. Li et al. [34] recently reported a replication, limiting their selection to those charts that consist of data sets with 10 or more observations. They also found that embellished charts are more enjoyable than plain ones.

Saket et al. [50] subsequently argued that although the study by Bateman et al. identified which visualization type is more enjoyable, it did not identify *what* makes one type of visualization more enjoyable than another. The researchers then suggested a model for measuring enjoyment and flow in visualizations. Their proposed model is built based on the Flow model of Csikszentimihayli [15] and Munzner's nested model [40] of visualization evaluation. More specifically, the model suggests that enjoyment in information visualization encompasses six different elements, to measure enjoyment evaluations must control as many of these elements as possible.

Challenge: Challenges in a visualization should match the skills of the user who is working with the visualization.

Focus: the user should be able to have a complete attention on the task at the hand.

Clarity: the user must understand exactly what the task's goals are.

Feedback: Visualization provides immediate feedback about the progress with he task at the hand.

Control: the user feels a complete control over the visualization itself and interaction techniques available to him/her.

Immersion: the user loses his/her sense of self and become "lost" in the process of working with a visualization.

In a subsequent study, Saket et al. measured the enjoyment and fun of two different node-link and node-link-group visualizations using their enjoyment model [51]. The study determined that participants found map-like visualizations more enjoyable than node-link visualizations.

In this small set of studies examining enjoyment of visualizations, the researchers used self-reporting methods such as Likert scale questions and interview questions to gain subjective information about participants' perceptions of a visualization's enjoyability. The common take-away was that embellished visualizations or ones with more pictorial representations tended to be more enjoyable. This is very initial research, however, and further work needs to be done to identify the particular aspects of a visualization that might make it more enjoyable and fun.

3. DISCUSSION

3.1 Debating the Merits

The importance of factors outside of the traditional usability and performance objectives for data visualization research is not a universally agreed upon view. Some researchers and practitioners have questioned whether issues such as memorability and enjoyment are important enough to merit further study.

In one notable example, Borkin et al. [9] published a study in 2013 that evaluated the memorability of a variety of real world visualizations. Visualization practitioner Stephen Few commented on this study in his blog, arguing that it simply assessed whether participants could remember a visualization's design, not its content, and that it did not reveal anything about what makes a visualization actually memorable [20]. His criticisms centered on a belief that visualizations cannot be read and understood in one second.

In 2015, Borkin et al. [8] conducted a more comprehensive study to move beyond memorability and investigate how visualizations are recognized and recalled. However, very soon after that in a newsletter entitled "Information Visualization Research as Pseudo-Science", Few again critiqued the previous work on memorability and more specifically this most recent study [21]. Few stated:

> *"Visualizations don't need to be designed for memorability – they need to be designed for comprehension. For most visualizations, the comprehension that they provide need only last until the decision that it informs is made. Usually, that is only a matter of seconds."*

Following Few's review, Ben Jones [30], a data visualization practitioner, responded to Few's newsletter. Jones stated:

> *"This statement [mentioned above by Stephen Few] helped me understand why Few and I disagree about memorability: we disagree about how data visualizations are used by groups of people. Simply put, I don't believe data visualizations are usually followed by decisions only a matter of seconds later. That may be how a robot or a computer algorithm would approach decision-making, but it's just not how groups of humans in organizations go about it."*

Later, in his own blog, Robert Kosara also responded to Few's article and advocated for investigating memorability in visualizations [31]. From his point of view, many cases exist where people would benefit from their audience remembering the visualizations (e.g., visualizations showing health related information).

We surmise that all visualization researchers consider a visualization's ability to accurately and effectively convey an understanding of the data it represents to be paramount. Debates such as those described above also make it clear that the importance of user experience-related goals is not as clear or universally agreed upon. Just how important is it for a visualization to be memorable, engaging, and/or fun? Moreover, do those attributes come at a corresponding cost? Research questions such as these addressing the value and merits of different characteristics arise in addition to more basic questions on how to simply measure and evaluate them. This article primarily has focused on the methodological issues involved in user experience metrics, but we felt it important to at least discuss the value issues and raise fundamental questions about the metrics' importance in the field as a whole.

3.2 Evaluation Methodology Challenges

As is evident from the studies reviewed earlier, a variety of evaluation methods have been employed to assess user experience metrics. One key differentiating factor appears to be whether subjective, often self-reported, measures are employed or more objective measures are gathered.

Memorability seems a little different than the other two prime goals we profiled. It easily lends itself to an objective measurement—Can a person remember a visualization or not? But even this question is not so simple. Many of the memorability studies tested whether a person could recall the imagery associated with a visualization, but not anything about the data being represented and its "story" or message. Aren't those things ultimately what the designer of a visualization wants their audience to take away, not just the visualization's appearance and style?

To assess engagement, many studies measured the time a person spent viewing or interacting with a visualization, which is clearly one form of engagement. But how about stronger immersion and a visualization's ability to foster people wanting to know more about the data and its domain? Facilitating deep(er) interest is another type of engagement that appears to be more difficult to objectively assess.

Finally, the sense of enjoyment and fun a visualization generates is typically measured through self-reported subjective methods such as interviews and Likert-style questions. Self-reported data often is questioned, however. Perhaps more objective measures such as counting smiles or laughs from people could be explored. It also seems clear that many forms of acknowledgment on social media such as page links, visits, likes, etc., can be correlated with stronger senses of engagement and enjoyment. These last metrics, e.g., counting likes, also highlight methodological issues of whether evaluations are carried out in controlled experimental settings or if the evaluation occurs in the field in a more natural setting and environment. Earlier work by Carpendale [14] and Lam et al. [33] highlight these and many more methodological issues that are equally important for user experience evaluation as for usability and performance evaluation.

3.3 Moving Forward

The previous subsections highlight some of the complex issues and challenges that are resident for this topic. Others surely exist as well, some relating back to the central questions about the importance of these experience-driven metrics. For example, a visualization can be memorable for many different reasons, including one such as being exceptionally bad. It is important not to group together all evaluations and metrics into very coarse categorizations. Surely, designers do not want their visualizations to be highly memorable because of their poor quality. We envision that user experience metrics should be assessed in coordination with other more traditional performance-driven measures in order to identify worthwhile and beneficial experience characteristics.

Another limitation of most of the studies we reviewed in this article is that they involve static visualizations. The importance of interaction in visualization is clear [60, 47, 18] and thus to focus only on static visuals simply appears to be too limiting. In fact, the interactive aspects of some visualizations (e.g., the earlier mentioned Baby Names Wizard) often strongly contribute to their high level of engagement and enjoyability. Of course, static visuals are easier to assess in controlled experiments—an interactive visualization allows varied exploration paths and hence comparisons between participants may be less reliable.

One interesting avenue for continued research is to understand the role that user interaction plays in making visualizations more engaging or enjoyable. Does adding a certain set of user interactions (e.g., zooming and panning, or linking and brushing) make a visualization more engaging? Does adding affordances to visualizations to help explain their functionality make them more appealing to users? Can one measure the contribution of user interaction versus visual representation to enjoyment and engagement? Are touch-based visualizations (e.g., tablet-based visualizations) more enjoyable and engaging than visualizations on desktops? We anticipate that user interaction plays a crucial role in making a visualization more engaging or enjoyable, but this remains to be formally studied.

Another future research direction is to investigate the impact of display size in which visualizations are shown on user experience. Do visualizations on large displays increase user engagement and enjoyment?

Additionally, it is absolutely crucial to consider the ultimate purpose or objective of a visualization when evaluating its user experience attributes. It is clear that many of the criticisms launched against assessing user experience metrics, such as those by Few [20, 21], are posed in light of visualizations used for analytical, exploratory purposes. This application is often considered first when reviewing visualizations. Only recently has the value of using visualizations for presentational and narrative purposes arisen in the visualization research community [53, 27, 32], even if this application is common outside research settings. We suspect that the user experience goals and metrics reviewed in this article play a more important role on visualizations used for narrative and storytelling purposes. Engagement, fun, and memorability are all well-accepted, desirable characteristics of successful presentations. Talented speakers deliver presentations high in these characteristics.

Furthermore, the entire context of use of a visualization definitely plays a strong role in the importance of these experience-related goals. For business- and work-related use where people interact with a visualization daily, a visualization's ability to effectively represent data and its information density likely will be more important than memorability and fun. Conversely, for entertainment contexts such as pass-by browsing of websites, billboards, or commercials, where catching a person's attention is vital, then user experience characteristics rise in importance.

Finally, the three main user experience goals reviewed in this article clearly are not the only experience-related goals that researchers might examine. Other characteristics of visualizations, such as their ability to be entertaining, provocative, surprising, motivating, and perhaps even exciting, all could be important toward a particular visualization being considered "successful." We chose to examine the three attributes discussed in this article because user studies exploring those attributes already had been published. Little, if any, research has explored other experience-related characteristics, so this is definitely an area for future opportunities and new work.

4. CONCLUSION

This paper discusses goals, metrics, and methods for evaluating data visualization techniques beyond traditional usability goals, such as efficiency and effectiveness. Performance-focused metrics such as time and accuracy are applicable in specific tasks and usage scenarios. Alternatively, methods focused on user experience goals are less frequently found in the visualization literature. This paper describes three categories of user experience goals: memorability, engagement, and enjoyment (or fun). Our discussion summarizes the study methodologies and metrics used to evaluate visualizations based on these three user experience goals. We summarize each of these by observing trends in current approaches, as well as highlight potential challenges and opportunities going forward. Our aim is to encourage visualization researchers to consider these user experience focused evaluation methodologies, where appropriate. Additionally, the specific challenges presented in this paper can lead to important research in evaluation methodologies for data visualization.

5. REFERENCES

[1] Baby Name Wizard - Name Voyager. http://www.babynamewizard.com/voyager.

[2] S. Attfield, G. Kazai, M. Lalmas, and B. Piwowarski. Towards a science of user engagement (position paper). In *WSDM workshop on user modelling for Web applications*, Feb. 2011.

[3] D. Baldauf, E. Burgard, and M. Wittmann. Time perception as a workload measure in simulated car driving. *Applied ergonomics*, 40(5):929–935, 2009.

[4] C. M. Barnum and L. A. Palmer. More than a feeling: understanding the desirability factor in user experience. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, pages 4703–4716. ACM, 2010.

[5] S. Bateman, R. L. Mandryk, C. Gutwin, A. Genest, D. McDine, and C. Brooks. Useful junk?: The effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2573–2582, New York, NY, USA, 2010. ACM.

[6] J. Benedek and T. Miner. Measuring desirability: New methods for evaluating desirability in a usability lab setting. In *Proceedings of Usability Professionals Association*, pages 8–12, July 2002.

[7] R. Borgo, A. Abdul-Rahman, F. Mohamed, P. W. Grant, I. Reppa, L. Floridi, and M. Chen. An empirical study on using visual embellishments in visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2759–2768, 2012.

[8] M. Borkin, Z. Bylinskii, N. Kim, C. Bainbridge, C. Yeh, D. Borkin, H. Pfister, and A. Oliva. Beyond memorability: Visualization recognition and recall. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):519–528, Jan 2016.

[9] M. Borkin, A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, H. Pfister, et al. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2306–2315, 2013.

[10] J. Boy, F. Detienne, and J.-D. Fekete. Storytelling in information visualizations: Does it engage users to explore data? In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 1449–1458, New York, NY, USA, 2015. ACM.

[11] P. B. Brandtzæg, A. Følstad, and J. Heim. Enjoyment: lessons from karasek. In *Funology*, pages 55–65. Springer, 2005.

[12] M. Brehmer and T. Munzner. A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, 2013.

[13] J. Brown, V. Lewis, and A. Monk. Memorability, word frequency and negative recognition. *The Quarterly Journal of Experimental Psychology*, 29(3):461–473, 1977.

[14] S. Carpendale. Evaluating information visualizations. In A. Kerren, J. T. Stasko, J.-D. Fekete, and C. North, editors, *Information Visualization*, pages 19–45. Springer-Verlag, Berlin, Heidelberg, 2008.

[15] M. Csikszentmihalyi. *Flow: The Psychology of Optimal Experience*. Harper Perennia, New York, 1990.

[16] W. A. Davis. A causal theory of enjoyment. *Mind*, 91(362):240–256, 1982.

[17] A. Dix. *Human-computer interaction*. Springer, 2009.

[18] N. Elmqvist, A. Vande Moere, H.-C. Jetter, D. Cernea, H. Reiterer, and T. J. Jankun-Kelly. Fluid interaction for information visualization. *Information Visualization*, 10(4):327–340, Oct. 2011.

[19] X. Feng, S. Chan, J. Brzezinski, and C. Nair. Measuring enjoyment of computer game play. In *Proceedings of AMCIS 2008*, 2008.

[20] S. Few. *Review of the Research Study "What Makes a Visualization Memorable?"*, 2013 (accessed February 3, 2016). https://www.perceptualedge.com/blog/?p=1770.

[21] S. Few. *Information Visualization Research as Pseudo-Science*, 2015 (accessed February 3, 2016). https://www.perceptualedge.com/articles/visual_business_intelligence/infovis_research_as_pseudo-science.pdf.

[22] B. Fredrickson. What good are positive emotions? *Review of General Psychology*, 3:300–3019, 1998.

[23] F.-L. Fu, R.-C. Su, and S.-C. Yu. Egameflow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education*, 52(1):101–112, Jan. 2009.

[24] S. Haroz, R. Kosara, and S. L. Franconeri. Isotype visualization: Working memory, performance, and engagement with pictographs. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 1191–1200, New York, NY, USA, 2015. ACM.

[25] L. Harrison, D. Skau, S. Franconeri, A. Lu, and R. Chang. Influencing visual judgment through affective priming. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 2949–2958, New York, NY, USA, 2013. ACM.

[26] J. Heer, F. van Ham, S. Carpendale, C. Weaver, and P. Isenberg. Creation and collaboration: Engaging new audiences for information visualization. In A. Kerren, J. T. Stasko, J.-D. Fekete, and C. North, editors, *Information Visualization*, pages 92–133. Springer-Verlag, Berlin, Heidelberg, 2008.

[27] J. Hullman and N. Diakopoulos. Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2231–2240, Dec. 2011.

[28] A. Isen. Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology*, 52(6):1122–1131, 1987.

[29] R. Jianu, A. Rusu, Y. Hu, and D. Taggart. How to display group information on node-link diagrams: an evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 20(11):1530–1541, 2014.

[30] B. Jones. *When Memorability Matters: Another Practitioner's View*, 2015 (accessed February 3, 2016). http://dataremixed.com/2015/12/when-memorability-matters-another-practitioners-view/.

[31] R. Kosara. *Memorability, Science, and The Value of Thinking Outside the Box*, 2015 (accessed February 3, 2016). https://eagereyes.org/blog/2015/memorability-science-and-the-value-of-thinking-outside-the-box.

[32] R. Kosara and J. Mackinlay. Storytelling: The Next Step for Visualization. *Computer*, 46(5):44–50, May 2013.

[33] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1520–1536, Sept. 2012.

[34] H. Li and N. Moacdieh. Is "chart junk" useful? an extended examination of visual embellishment. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 58, pages 1516–1520. SAGE Publications, 2014.

[35] N. Mahyar, S.-H. Kim, and B. C. Kwon. Towards a taxonomy for evaluating user engagement in information visualization. *Workshop on Personal Visualization: Exploring Everyday Life*, 2015.

[36] T. W. Malone. Toward a theory of intrinsically motivating instruction. *Cognitive science*, 5(4):333–369, 1981.

[37] T. W. Malone. Heuristics for designing enjoyable user interfaces: Lessons from computer games. In *Proceedings of the 1982 Conference on Human Factors in Computing Systems, CHI '82*, pages 63–68, New York, NY, USA, 1982. ACM.

[38] K. Marriott, H. Purchase, M. Wybrow, and C. Goncu. Memorability of visual features in network diagrams. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2477–2485, 2012.

[39] T. Merčun. Evaluation of information visualization techniques: Analysing user experience with reaction cards. In *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*, BELIV '14, pages 103–109, New York, NY, USA, 2014. ACM.

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

[41] D. A. Norman. *Emotional design: Why we love (or hate) everyday things*. Basic books, 2004.

[42] D. A. Norman. *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[43] H. L. O'Brien and E. G. Toms. What is user engagement? a conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6):938–955, 2008.

[44] H. L. O'Brien and E. G. Toms. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, 61(1):50–69, 2010.

[45] F. Pachet and A. R. Addessi. When children reflect on their own playing style: Experiments with continuator and children. *Comput. Entertain.*, 2(1):14–14, 2004.

[46] R. Parncutt and G. E. McPherson. *The Science & Psychology of Music Performance: Creative Strategies for Teaching and Learning Book*. Oxford University Press, 2009.

[47] W. A. Pike, J. Stasko, R. Chang, and T. A. O'Connell. The science of interaction. *Information Visualization*, 8(4):263–274, Dec. 2009.

[48] K. Rathunde and M. Csikszentmihalyi. Middle school students's motivation and quality of experience: A comparison of montessori and traditional school environments. *American Journal of Education*, 111:341–371, 2005.

[49] J. C. Read, S. MacFarlane, and C. Casey. Endurability, engagement and expectations: Measuring children's fun. In *Interaction design and children*, volume 2, pages 1–23. Shaker Publishing Eindhoven, 2002.

[50] B. Saket, C. Scheidegger, and S. Kobourov. Towards understanding enjoyment and flow in information visualization (short paper). In *EuroVis 2015 Conference*, 2015.

[51] B. Saket, C. Scheidegger, and S. G. Kobourov. Comparing node-link and node-link-group visualizations from an enjoyment perspective. *Computer Graphics Forum*, 35(3):41–50, 2016.

[52] B. Saket, C. Scheidegger, S. G. Kobourov, and K. Börner. Map-based visualizations increase recall accuracy of data. *Computer Graphics Forum*, 34(3):441–450, 2015.

[53] E. Segel and J. Heer. Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139 –1148, Dec. 2010.

[54] H. Sharp, Y. Rogers, and J. Preece. *Interaction Design: Beyond Human Computer Interaction*. John Wiley & Sons, 2007.

[55] B. Shneiderman. Designing for fun: how can we design user interfaces to be more fun? *interactions*, 11(5):48–50, 2004.

[56] P. Sweetser and P. Wyeth. Gameflow: A model for evaluating player enjoyment in games. *Computers in Entertainment*, 3(3),

July 2005.

[57] M. Vass, J. M. Carroll, and C. A. Shaffer. Supporting creativity in problem solving environments. In *Proceedings of the 4th Conference on Creativity & Cognition*, pages 31–37, New York, NY, USA, 2002. ACM.

[58] M. Wattenberg and J. Kriss. Designing for social data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 12(4):549–557, July 2006.

[59] L. Xie, A. N. Antle, and N. Motamedi. Are tangibles more fun?: Comparing children's enjoyment and engagement using physical, graphical and tangible user interfaces. In *Proceedings of the 2nd international conference on Tangible and embedded interaction*, pages 191–198. ACM, 2008.

[60] J. S. Yi, Y. a. Kang, J. Stasko, and J. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1224–1231, Nov. 2007.