Chapter 4
Improving Evaluation Using Visualization Decision-Making Models: A Practical Guide

Melanie Bancilhon, Lace Padilla, and Alvitta Ottley

Abstract In visualization research, evaluation is a crucial step to assess the impact of visualization on decision-making. Existing work often gauges how good a visualization is by measuring its ability to induce accurate and fast judgment. While those measures provide some insight into the efficacy of a graph, underlying cognitive processes responsible for reasoning and judgment are often overlooked when they can have significant implications for visualization recommendation. Cognitive processes do not need to be a black box. There exists multiple models that describe decision processes, such as theories from behavioral economics and cognitive science. In this chapter, we compare and contrast different models and advocate for the inclusion of cognitive models for visualization evaluation in the context of decision-making. The goal of this work is to show visualization researchers the advantages of adopting a more mechanistic approach to evaluation at the intersection of visualization and cognitive science.

4.1 Introduction

We make decisions based on data every day, ranging from trivial to complex. Such choices could include when to leave the house to catch the bus, take an umbrella given the chance of rain, or invest in the stock market given the historical trends. In many instances, charts and graphs have become an integral part of our decision-making process. Visualization research has provided valuable insight into perceptual science and has led to the amelioration of chart design and visualization recommendations. Charts frequently appear in information communication, data analysis, sensitization campaigns, and even medical diagnostics and can significantly impact

M. Bancilhon (✉) · A. Ottley
Washington University in St. Louis, St. Louis, MO, USA
e-mail: mbancilhon@wustl.edu; alvitta@wustl.edu

L. Padilla
Northeastern University, Boston, USA
e-mail: lace.padilla@ucmerced.edu

D. Albers Szafir et al. (eds.), *Visualization Psychology*,
https://doi.org/10.1007/978-3-031-34738-2_4

people's lives. But all charts are not equal. When a new graph or chart is designed, it is essential to conduct an evaluation under realistic decision-making conditions to understand and foresee its effect on real-life decisions.

However, it can be hard to know if an evaluation is close enough to natural decision-making conditions to provide meaningful insights into the efficacy of a visualization. One way to conduct rigorously valid evaluations is to understand and simulate the underlying mental mechanisms at work when a viewer completes the real-world task. Fortunately, cognitive scientists have extensively studied cognitive mechanisms responsible for interpreting and misinterpreting visual designs under different modes of reasoning. For example, *dual-process theory* posits that there exists two types of decisions operating under distinct cognitive processes: intuitive (Type 1) and strategic (Type 2) decisions, which require significantly more effort than Type 1 [33]. In this chapter, we dive into multiple prominent perspectives of decision-making. We discuss how the researchers can apply frameworks and models pertaining to visualization design and evaluation in the context of decision-making. We propose that dual-process cognitive models are some of the most useful and easily applied for visualization research. This chapter will be helpful for designers and visualization researchers looking to adopt a more granular approach to decision-making and conduct holistic evaluations for better visualization recommendations.

4.1.1 Evaluation Methods for Decision-Making

Research on visualization evaluation is vast and varied [17, 43], with high tendencies toward evaluating visualization based on speed and accuracy in perceptual judgments [64]. A relatively small number of studies have focused on evaluating people's visualization-aided decisions. Researchers have investigated how visualizations impact attitudes toward risk and hypothetical decisions [22, 62]. For example, Ruiz et al. [62] conducted a study where they asked at-risk patients to decide whether they would opt for screening based on hypothetical risk information about a disease [62]. They found that people are more risk-averse when presented with icon arrays. Kay et al. [37] evaluated how well different visualizations communicate the uncertainty of transit data by asking participants to estimate the likeliness of bus arrival times on a scale of 0 to 100 [37].

In traditional visualization empirical studies, visualizations are often evaluated by their ability to prompt accurate and fast responses in behavioral tasks, that may or may not involve making a decision. While it is common to extrapolate the appropriateness of visualizations for decision-making through these performance-based measures, there are less attempts to evaluate visualization designs based on the quality of the decisions they elicit [51]. Empirical evaluations of visualization are generally challenging [9, 17, 56]. Thus, one possible reason for the lack of evaluations with decision-making is that it is generally more straightforward to gauge effectiveness via the speed and accuracy of perceptual judgments. Consider, for example, the chart shown in Fig. 4.1, which shows a given person's chance

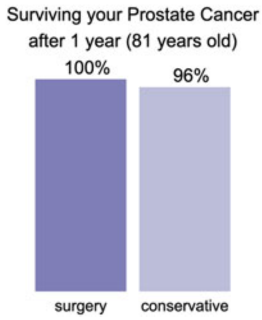

Fig. 4.1 A bar chart comparing the survival rates after one year of surgery versus conservative management for a 80-year-old prostate cancer patient [23]

of surviving prostate cancer after one year if they choose to have surgery (e.g., radical prostatectomy) compared to conservative treatments (e.g., watchful waiting). One could evaluate this chart based on how well it facilitates fast and accurate comparisons of the two quantities, or based on the responses from semi-structured interviews with prostate cancer patients [23]. Experiment protocols like these are more straightforward than those that measure decisions because it is feasible to define a ground truth or expected behavior for the analysis of study findings.

In practice, we often use performance-based findings to inform the selection of visualization designs, implying that accurate decoding likely leads to better and more informed decisions. Based on our current understanding of perceptual judgments, the bar chart in Fig. 4.1 uses position for data encoding, and therefore is ideal for comparing quantities and seeing small differences [12, 13]. However, one could reasonably assert that the difference between the survival rates for surgery (100%) and conservative treatment (96%) is statistically insignificant, but the bar chart might inadvertently emphasize a potentially minor disparity. Existing studies show that the ideal visualization depends on the task. For example, the superior representation for magnitude estimation might not be optimal for part-to-whole judgments [20, 65, 66]. Some researchers have used simulations to observe the direct impact of visualization design on decisions. In one study by Bancilhon et al. [4], participants played a lottery game and chose to either enter the lottery or receive guaranteed monetary gains based on five standard visualization designs. They analyzed the quality of the decisions based on economic optimality and found that people made significantly more risk-seeking decisions with circle and triangle charts [4] (see Sect. 4.3.1.2).

Decision-making is complex and multifactorial. In addition to the graph's appropriateness, a patient's decision to have surgery (or not) will depend on various factors including illness severity age, commodities, and personal finances. People are also prone to various cognitive biases [16], and individual differences in personality and cognitive abilities may also influence usability and choice [40, 53]. At a fundamental level, the decision-maker's perspective drives the decision, and the typical approach of defining a ground truth in an evaluation is non-trivial. Despite this challenge, other fields have demonstrated success in modeling and predicting, and reasoning about how people make decisions [33, 35, 55, 57]. We argue that for visualization to be a practical tool for supporting decision-making, we need to understand the underlying cognitive processes behind decision-making and adopt a unifying cross-discipline framework to evaluate visualization in this context.

To aid this discussion, we adapt Balleine's definition of decision [3]:

> A *decision* is a choice between competing courses of actions [3].

4.2 The Science of Making Decisions

Decisions are governed by complex systems of reasoning that scholars have studied for decades. Researchers in the visualization community have pursued two dominant approaches to study decision-making under risk. The first provides a detailed and quantifiable view of decision-making. It assumes that humans make decisions rationally by weighing the risk and expected outcome of different prospects, two factors that can be measured and modeled. The second posits that many factors can influence decision-making. It proposes that humans make both intuitive (Type 1) and strategic (Type 2) decisions and that decision-makers usually default to using intuition. These two distinct types of decisions operate under a *dual-process system*. To improve visualization research in the context of decision-making, it is crucial to understand the meaning and implications of decision-making under both umbrellas. We structure this chapter around two prevalent approaches: *The Utility-Optimal Perspective* and *The Dual-Process Perspective*.

4.3 The Utility-Optimal Perspective

Behavioral economists have long studied how people make choices under risk by investigating prospects or gambling scenarios. A prospect is a contract:

$$[(x_1, p_1), (x_2, p_2), \ldots, (x_n, p_n)], \tag{4.1}$$

which yields x_i with probability p_i, where $\sum_{i=1}^{n} p_i = 1$ [35]. Prospects provide a simple model for understanding risky decisions. The classical method for evaluating a gamble is through assessing its expected value. The expected value of a prospect is the sum of the outcomes where the probabilities weigh each value:

$$ev = \sum_{i=1}^{n} p_i x_i. \tag{4.2}$$

Consider the gambling scenario from Kahneman and Tversky's book [35]:

Which do you prefer?
Option A: 50% chance to win \$1000, 50% chance to win \$0
Option B: \$450 for sure

The expected value of option A is 500 ($.5 \times 1000 + .5 \times 0$) and the expected value of option B is 450 (1×450). A *rational* decision-maker would then choose option A over option B. However, most people would choose the sure payment of \$450. This example highlights the perhaps obvious conjecture that humans are not always rational [35].

Expected Utility Theory (EUT) is one of the foundational theories of decision-making and has served for many years as both a model describing economic behavior [21] and a rational choice model [38]. In particular, it states that people make choices based on their *utility*—the psychological values of the outcomes. For instance, if a person prefers an apple over a banana, it stands to reason that they would prefer a 5% chance of winning an apple over a 5% chance of winning a banana. Using EUT, we can assess the overall utility of a gamble:

$$EU = \sum_{i=1}^{n} p_i u(x_i), \tag{4.3}$$

where the function u assigns utility to an outcome. We sum the utilities u of the outcomes x_i weighted by their probabilities p_i. This model has its limitations. It also assumes that humans are consistent and primarily decide on prospects based on their utility [35, 69]. Nevertheless, EUT provides a standardized tool for researchers to evaluate peoples' behavior when choosing among risky options and is the foundation for the other dominant theory in behavioral economics, *Prospect theory* [35].

Unlike EUT, prospect theory embraces the human factors present in decision-making. Kahneman and Tversky [35] are the pioneer contributors to this knowledge on bias in decision-making. For example, in their early work, they found that 72 out of 100 experiment participants favored the option of getting \$5000 with a probability of 0.001 (e.g., a small probability event) over the prospect of getting \$5 for sure [35]. Both options have the same expected value, yet most participants favored the probability associated with getting \$5000. In its simplest form, we can represent the equation for prospect theory as

$$V = \sum_{i=1}^{n} \pi(p_i)\upsilon(x_i), \tag{4.4}$$

where the function υ assigns value to an outcome and the function π is a probability weighing function that encodes the idea that people are likely to overreact to small probabilities and underreact to large probability events. In summary, prospect theory stipulates that (1) people tend to favor the option of getting a large gain with a small probability over getting a small gain with certainty and (2) people tend to prefer a small loss with certainty over a large loss with tiny probability.

4.3.1 Using Utility-Optimality to Evaluate visualizations

Visualization researchers have leveraged utility-optimal theories to investigate how visualization impacts decisions under risk. By approaching decision-making from this angle, they create an environment where choices have weights, and their evaluation considers the utility-optimal option. We highlight two empirical studies from the visualization community and examine their experimental design, methodology, and research questions. We will begin with a recent publication investigating the impact of uncertainty visualization design by simulating a fantasy football scenario.

4.3.1.1 A Fantasy Football Study

Kale et al. [36] leveraged utility-optimal theories to observe effect size judgments and decision-making with the four uncertainty visualizations. They used a fantasy football game to elicit decisions under risk. Participants were shown the number of points scored by a certain team with and without the addition of a new player. First, they asked participants to estimate a measure of effect size by asking the following question: *"How many times out of 100 do you estimate that your team would score more points with the new player than without the new player?"*. They also asked participants to make binary decisions indicating whether they would *Pay for the new player* or *Keep their team without the new player*. On each trial, the participant's goal was to win an award worth $3.17M, and they could pay $1M to add a player to their team if they thought the new player improved their chances of winning enough to be worth the cost.

They tested four uncertainty visualizations: 95% containment intervals, hypothetical outcome plots (HOPs), density plots, and quantile dot plots, each with and without means added. They found that while adding means to quantile dot plots produced significantly more utility-optimal decisions at low variance, it had no reliable effect on bias in magnitude estimation. Similarly, adding means to HOPs caused significantly more bias in magnitude estimation across both low and high variance but had no reliable effect on decisions. By evaluating uncertainty visualizations using utility-optimality, Kale et al. [36] observed a decoupling of performance across tasks, where the visualization designs that support the least biased effect size estimation do not support the best decision-making and vice versa. The authors attribute this inconsistency to the reliance on different heuristics across the two different tasks, consistent with Kahneman and Tversky's theory [35]. This finding highlights the value of leveraging utility-optimal theories when studying visualization for decision-making.

4.3.1.2 A Classic Lottery Game

Many studies that leveraged utility-optimal decision-making theories employed tasks with hypothetical gains and losses (e.g., [10, 31, 36, 49]). However, it is unclear if people make the same risk judgments when gains and losses do not tangibly affect them. To evaluate visualization decision-making with greater *ecological validity* (i.e., more closely matching real-world conditions), Bancilhon et al. [4] created a gambling game that immersed participants in an environment where their actions impacted the bonus payments they received. The experiment investigated the effect of five charts on decision-making. Replicating the experiment design of prior work in the economic decision-making domain [8], the researchers presented participants with two-outcome lotteries: take the sure gain or gamble at a risk. The experiment employed a point system for payoff quantities where 1 point equaled $0.01. The probabilities, p_i, were drawn from the set $P = \{.05, .1, .25, .5, .75, .9, .95\}$ and the outcomes x_1 and x_2 ranged from 0 to 150 points ($0 to $1.50).

Figure 4.2a shows an example of the lottery sheet used in the study. At the end of the experiment, the game randomly selected one row from each of the 25 lottery sheets that they saw, and the participant's choice in that row determined their bonus. If the participant chose the sure payout in the selected row, their bonus increased by that amount. If they opted to enter the lottery, the game simulated the lottery to determine the payment, with the potential gains and the probabilities as parameters.

Overall, the findings from the study [4] validate that we can use utility-optimal theories to evaluate visualization designs, and that the latter can influence gambling behavior. They had three major findings. First, the *icon array* was most likely to elicit risk neutrality and is, therefore, the most effective design for decision-making. Second, they found that participants who saw a *bar* chart exhibited behavior that was slightly risk-averse, mirroring behavior in the control text-only group. Third, the *triangle* chart and *circle* chart elicited risk-seeking behavior with the greatest deviation from risk neutrality. It is important to note that these findings are in line with the magnitude estimation from the prior literature [13] that shows that proportion estimates with *bar* charts are more accurate than with *triangle* and *circle* charts.

4.3.2 Outlook on Using Utility-Optimal Theories for Visualization Evaluation

Although we only highlighted a few studies in this section, it is essential to note that other researchers have also examined decision-making with visualization using a similar framework (e.g., [10, 26, 31, 49, 71]). For example, Padilla et al. [49] conducted a scenario where participants made resource allocation judgments by comparing the cost of sending cold-weather aid to alpaca farmers in Peru who were at risk of losing their livestock due to cold temperatures and the expected

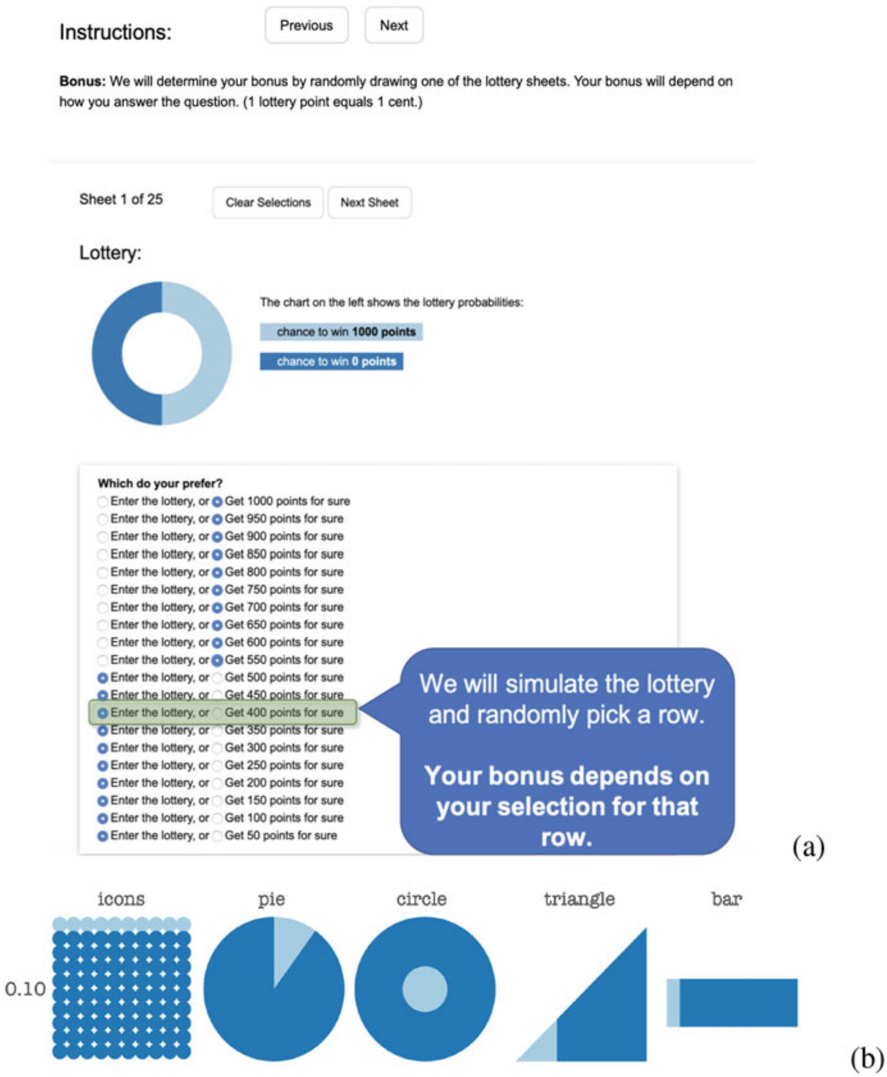

Fig. 4.2 The charts and lottery sheet used in the study by Bancilhon et al. [4]. Participants played a gambling game in which their choices determined their bonuses

value of the penalty for not sending aid, resulting in the deaths of alpacas (see also, [10, 31]). Perhaps most importantly, for visualization evaluation, the utility-optimal perspective provides a tractable approach to quantifying and modeling decision-making under risk. In both Kale et al.'s and Bancilhon et al.'s studies [4, 36], the researchers leveraged the framework to isolate the effect of visualization design. In some cases, their results suggest that using visualizations might help to reduce biases and guide people towards utility-optimality [4].

It is typical for researchers to design games or simulations to observe people's decisions in action. In many cases, it is difficult, if not impossible, to test the impact of visualizations on decisions in real life as it may give rise to safety, health, and ethical issues. For example, it might be unsafe and unethical for a gambling game to test the effect of visualizations that communicate information about a severe health condition that a participant has or a natural disaster affecting the participant at the time of the study. The utility-optimal framework using the situational scenarios in the two studies [4] and [36] provides a good test bed for evaluating visualizations for decision-making. In order to apply this framework to behavioral studies, there needs to be a cost associated with each course of action. The utility-optimal decision should be defined as the one where prospective gains are maximized and losses are minimized. By quantifying user choices and comparing them to the utility-optimal decision, we can infer the risk behavior elicited by the visualization design. It is important to take into account people's patterns of risk behavior since humans do not normally default to risk neutrality regardless of the type of representation used. By providing an incentive to decision-makers, such an experiment design can more closely mimic real-life choices over hypothetical decision scenarios.

While Bancilhon et al. [4] have shown that the visualizations that lead to better accuracy also induce more optimal decisions, Kale et al. [36] have shown that the visualization designs that lead to the least bias did not lead to the most optimal decisions and vice versa. First, this shows that task and visualization choice matter in evaluation. Second, it raises an important question: how do we define the best visualization when accuracy and utility-optimal decisions are inconsistent? In Kale et al.'s study [36], one approach to determine the best uncertainty visualization would be to pick the one with the best compromise between high accuracy and optimal decision-making. Huang et al. [27] have developed a model of visualization efficacy that includes speed, accuracy, and cognitive load, which is often overlooked. One way forward could be to refine this model to include decision-making. Another approach would be to simply not attempt to choose a single best visualization for reasoning about uncertainty. Kale et al. [36] have shown that different visualizations are best for different tasks. There needs to be a common recognition in the visualization community that a one-size fits all approach could be obsolete.

Furthermore, using utility-optimality for visualization evaluation raises another crucial question: how do we define the *best* decisions? Some would argue that rationality should be the golden standard since it maximizes the potential outcome. Bancilhon et al. [4] question whether or not that should be the case. If the goal is rationality, their findings suggest that the icon array was the most likely to elicit risk-neutral choices. However, since people make decisions according to their personal inclination to risk, there might be a cost in attempting to steer them toward utility-optimality. Perhaps an ideal visualization should support the users in making a decision based on their individual risk behaviors.

In the next section, we examine a different perspective on decision-making, positing that humans default to intuitive reasoning when making decisions. We discuss working memory as a metric for usability in visualization decision-making (Fig. 4.3).

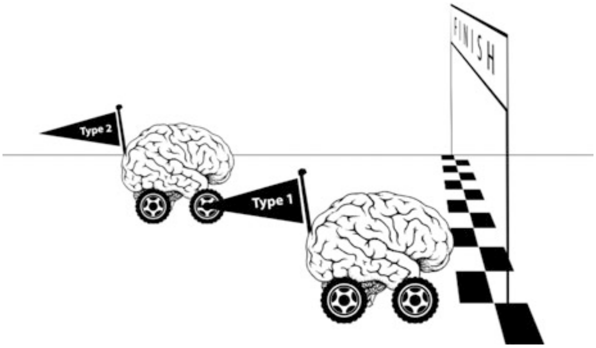

Fig. 4.3 An illustration of Type 1 and Type 2 reasoning as conceptualized by Tversky and Kahneman [33]. Type 1, our intuitive system, is at the forefront of decision processes, while Type 2, our analytic system, operates secondarily

Working memory consists of various mental components that can hold a limited amount of transformable information for a finite period [14]. In visualization research, working memory is commonly associated with mental effort [47]. Note that there is an ongoing debate on the definition of working memory [14]

4.4 The Dual-Process Perspective

In addition to the biases associated with gains and losses (e.g., prospect theory), many other cognitive biases are involved when making decisions under risk. One perspective that describes a large body of biases proposes that people rely on quantitative reasoning and gist-based intuition—two systems that operate in parallel [33].

Daniel Kahneman published a book entitled *Thinking Fast and Slow*, where he summarized decades of research on a dual-system of decision-making [33]. In his earlier work, his collaborators and himself differentiated between two types of processing systems, termed *System 1* (or intuition) and *System 2* (or reasoning) [32] (later termed Type 1 and Type 2). Type 1 processing guides our intuition and recognition patterns, which occur automatically without effort. In contrast, Type 2 processing is responsible for analytical thinking and requires directed effort to use [33].

Dual-Process Theory introduces a reasoning model that formalizes the differences between Type 1 and Type 2 and their impact on decision-making [34, 67]. Proponents of Dual-Process Theory posit that most decisions stem from intuitive thinking rather than rational and calculated thinking [33]. Type 1 reasoning involves fast and intuitive thinking, while Type 2 is a slow and analytical method of thinking.

Scholars propose that Type 2 processing uses significant working memory, while Type 1 only uses negligible working memory [18]. Using this definition, the researchers can determine when people are using Type 2 processing by identifying when people show an increase in working memory demand. Visualization researchers have demonstrated how to measure an increase in working memory demand using pupillometry (e.g., dilation of pupils [47]), dual-tasking (e.g., doing two tasks a once [11, 47]), individual differences measures (e.g., working with participants with high- and low-working memory capacity [10]), the NASA-TLX (e.g., self-reported work-load [10]), and electroencephalography (e.g., neuroimaging [1]). Type 1 is at the forefront of cognitive processes, and it often requires significant effort to switch from Type 1 to Type 2 in order to avoid cognitive biases and misleading heuristics. Despite utilizing different strategies, dual-process theories propose that the processes do not necessarily occur in separate cognitive or neurological systems [19].

Other frameworks have adapted the general dual-process perspective as well. Notably, Reyna and Brainerd introduced *Fuzzy Trace Theory* (FTT) [58]. The theory posits that people form two types of mental representations from information: *Gist* and *Verbatim* representations. A verbatim representation is a detailed representation of an event that often comprises precise numbers and facts. Gist representation, on the contrary, is vague and high-level and captures the essential meaning of information. FTT asserts that people make decisions by extracting meaning from verbatim input to make a gist-based judgment. According to Reyna and Brainerd [58], the human memory contains various reasoning-relevant information, ranging from preserving the exact form of input or only retaining abstract representations. People operate somewhere between the highest level of gist and the highest level of verbatim, on a gist-to-verbatim continuum [58]. Typically, humans rely on the least precise gist representation necessary to make a decision, and this characteristic is generally referred to as "fuzzy processing preference" [58].

Although there is a long history of theories on dual-processes, the high-level ideas are similar. They assert that there are two kinds of reasoning. One is implicit, intuitive, and unconscious, and the other is explicit, conscious, and slow. For simplicity, we will refer to this general class of theories as *Dual-Process* theories.

4.4.1 Dual-Process in Decision-Making

Fuzzy Trace Theory states that people make decisions by extracting meaning from verbatim input to make a gist-based judgment. Because precision is often associated with accuracy, many believe that quantitative reasoning is superior to qualitative reasoning. However, in some cases, fuzzy representation of information does not affect reasoning accuracy [60]. Reyna and Lloyd [59] have shown that experts in the medical field tend to engage more in gist-based decision-making than novices. Tversky and Kahneman made the argument that intuition is a synonym for recognition [33]. Experts recognize familiar situations and can therefore make fast and accurate decisions even when they are complex.

Although Type 1 has been proven to be efficient [59, 60], it is also more susceptible to false first impressions and framing effects [33]. Consider the following question:

A bat and ball cost \$1.10. The bat costs \$1 more than the ball. How much does the ball cost?

More than 50% of students at Harvard, Princeton, and the Massachusetts Institute of Technology routinely gave the incorrect answer, insisting the ball costs 10 cents [33].[1] Type 1 is at the forefront of cognitive processes, and in order to obtain the correct answer, a switch from Type 1 to Type 2 is required to overcome cognitive biases.

Before the acknowledgement of the role of Type 1, many believed that Type 2 was solely in charge of decision-making operations. Expected Utility Theory posits that people make decisions rationally, using Type 2 to compute the utility of events. The recognition of dual modes of reasoning lead to the development of prospect theory [35] (see Sect. 4.3) and revolutionized decision-making research.

4.4.2 Dual-Processes and Visualization Evaluation

In the medical field, researchers have investigated the impact of visualization design on gist reasoning. For example, Feldman et al.'s first goal [20] was to investigate which graphical formats induced the most accurate perception of quantitative information in patients making treatment decisions. Second, they inquired about the formats that facilitate processing. The authors highlight the importance of ease of processing, especially when the patient feels overwhelmed by the diagnostic. They conducted an experiment to test the performance of variations of 6 visualization formats. Participants had to minimize how long the visualizations appeared on the screen while remaining accurate when answering questions about the charts. They were shown two quantities and were asked to make a gist judgment by choosing the one that showed the larger chance of survival or the smaller chance of side effects. They were also asked to make a verbatim judgment by determining the size of the difference.

In this study, Feldman et al. [20] used response time as a proxy for ease of information processing. Their results suggest that systematic ovals, which encode data in a natural frequency format, are likely the format that represents the best compromise for accurate processing of both gist and detailed information while also demanding relatively little effort. Similarly, Hawley et al. [24] conducted an experiment investigating gist and verbatim reasoning through similar comparison

[1] The correct answer to this problem is that the ball costs 5 cents and the bat costs –at a dollar more– \$1.05 for a grand total of \$1.10.

and estimation tasks. They found that viewing a pictograph was associated with adequate levels of both gist and verbatim knowledge and that superior medical treatment choices were made in both cases.

In their work, Feldman et al. [20] question the overall effectiveness of vertical bars with scales, which was the best visualization for gist reasoning. The authors state that many patients demand detail-level information, and they defined the best visualization as the one that is effective in eliciting both types of reasoning. While this prior work gives evidence that charts using natural frequency encoding perform better under both gist and verbatim reasoning in comparison tasks, further research is required to examine whether the findings are generalizable to other tasks.

4.4.3 Outlook on Using the Dual-Processing Approach for Visualization Evaluation

While the Expected Utility Framework provides a method to mathematically model decisions, the Dual-Process framework is not straightforward. Feldman et al. [20] and Hawley et al. [24] have studied how visualization affects Type 1 and Type 2 reasoning in a comparison task. Note that it is possible for both processes to be used to make a decision. In their respective work, they posit that a magnitude estimation task brings about Type 2 reasoning, whereas asking the participant to make a comparison choice triggers Type 1 reasoning. If we apply this inference to Bancilhon et al.'s lottery game study [4] in Sect. 4.3.1.2, their results are consistent with Feldman et al.'s work [20] since the icon array outperforms the other visualizations in the decision task. Considering Kale et al.'s fantasy football study [36] in Sect. 4.3.1.1, which observed a magnitude estimation task and a decision task, it is possible that the selected visualizations have different effects under Type 1 and Type 2 reasoning.

However, our conclusions are solely based on the assumption that the tasks used actually elicit two distinct types of reasoning. To further research in this area, we need to answer the following research questions, which are core to understanding the role of visualization in decision-making:

- How does the mode of reasoning influence decision-making when using visualizations?
- Can different visualizations elicit different modes of reasoning?

It is crucial to understand how people make decisions from visualizations. Understanding whether a visual encoding facilitates gist or verbatim reasoning can have enormous implications for visualization designers. By expanding our knowledge in this area, we can tailor visualizations to our audience or a specific problem area. Bridging the gap between how psychologists and visualization researchers reason about decision-making can revolutionize how we evaluate and design visualizations.

Such knowledge can have massive implications for visualization designers. For example, visualizations can be tailored and personalized to a specific problem area or level of audience expertise. Some visualizations are only seen for a short time so we need a quick way of displaying information so that people get the gist of it. Moreover, some people might be more prone to gisting and others to probabilistic reasoning. Factors such as numeracy and spatial ability likely play a role.

Further investigations are needed to understand *how* people reason under this dual mode and how it affects their decisions. In the following sections, we examine cognitive models of decision-making with visualization and advocate for their integration into visualization research to deepen our understanding of decision-making processes with different charts.

4.5 Cognitive Models of Decision-Making with Visualization

Cognitive models are an integration of approaches and can be illustrated as process diagrams that conceptualize their mechanisms processes. By applying a cognitive model to a problem, a visualization researcher can better understand, model, or even evaluate the interaction between the user and the visual design at a cognitive level of analysis, as opposed to strictly behavioral. Before diving into the integration of a dual-process approach into decision-making research with visualization, we must first understand how the mind perceives and understands visualization. Pinker [55] proposed a cognitive model depicting the distinction between two mechanisms in graph comprehension: bottom-up and top-down mechanisms [55].

Bottom-up processing is when the mind is directly influenced by a visual stimulus which is utilized to construct a visual description.
Top-down processing is based on the viewer's goals, experiences, and other individual differences.

Prior knowledge about the graph is then retrieved from long-term memory in the form of an established graph schema. It is essential to point out that with familiar charts, the visual schema will be retrieved from memory faster and more efficiently, facilitating Type 1 reasoning [48]. This *match process* also occurs when visual properties are altered. The viewer then retrieves the graph schema that is the most similar to the visual array. When a graph schema is retrieved, the viewer uses the information from the graph schema to interpret the visualization. Bottom-up attention is influenced by saliency in the visualization design. Features that attract bottom-up attention are color, edges, lines, and foreground information.

Graph schema is memorized graphic conventions [55].

When external factors impact knowledge retrieval, the viewer is considered to be taking a top-down approach. Top-down attention is based on the viewer's goals, experiences, and other individual differences. There are other factors that can affect visualization comprehension, such as the nature of the task. Viewers may need to transform their mental representation of the visualization based on their task or conceptual questions, and working memory plays are central role in the process (Fig. 4.4).

4.5.1 Padilla's Dual-Process Model and the Importance of Working Memory

Padilla et al. [48] devised a model that combines theories of visualization comprehension, decision-making, and working memory. The motivation for this work is the lack of formalization of research from different fields, making it difficult for scientists to integrate cross-domain findings. The authors explored a cognitive model of decision-making with visualizations and provide practical recommendations for visualization designers.

In the previous section, we defined two types of graph comprehension mechanisms: bottom-up and top-down. The understanding of these two mechanisms is crucial in the understanding of Padilla's Dual-Process Model, with the addition of working memory, which are the mental processes associated with effort [48].

Padilla et al. [48] assert that working memory plays an important role in decision-making, but it is often overlooked by visualization researchers as an evaluation factor. Before diving into how working memory is involved in the dual reasoning system, let's look at some of its properties. It is important to note that working memory capacity is limited [42, 63]. Working memory also increases with task difficulty and diminishes over time. Researchers such as Cowan et al. [15] suggest that our ability to store information begins to decay after approximately 5–10 seconds, depending on factors such as the task, type of information, and individual differences in working memory capacity. One property of working memory capacity that is relevant to dual-process theory is that it limits the amount of attention we can allocate to task-relevant information [48].

Padilla et al.'s model [48] suggests that when we deliberately employ working memory in our decision-making process, we can make slower and more strategic but cognitively demanding decisions with visualizations. In other words, working

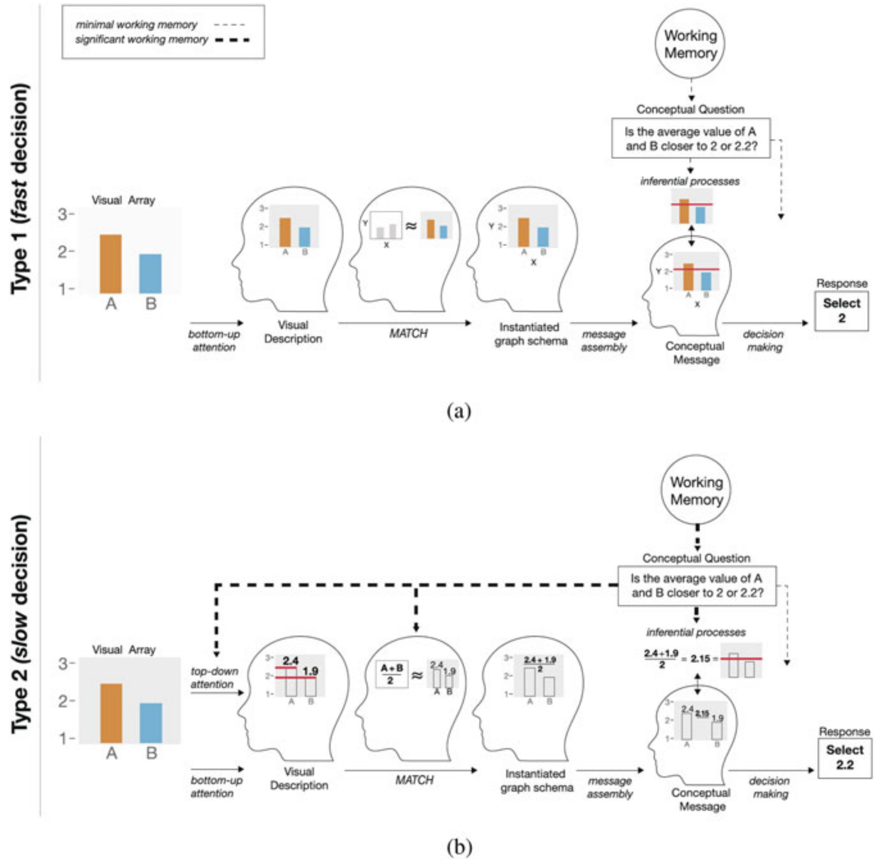

Fig. 4.4 An illustrative example of Type 1 versus Type 2 decision as characterized by Padilla et al.'s model [48]. (**a**) An example of a Type 1 decision process in which the viewer is tasked with computing the average of the two bars in the graph. A Type 1 approach might make a quick guess of the middle point between the two bars. (**b**) An illustration of a Type 2 decision process. The task is the same as subfigure (**a**) above. In this example, the viewer takes a slower approach and estimates the length of each bar. They then compute the average of the two values $\frac{2.4+1.9}{2}$. Type 2 activates working memory and can lead to a more effortful but precise estimate if done correctly

memory is what we use to switch from Type 1 reasoning (requiring nominal working memory) to Type 2 (requiring significant working memory). As described in the previous section, both Type 1 and Type 2 reasoning can be used to complete the decision step. Differences in working memory capacity can influence judgments and consequently decision-making. Lohse [41] found that when participants made judgments about budget allocation using profit charts, individuals with less working memory capacity performed equally well compared to those with more working memory capacity when they only made decisions about three regions (easier task). However, when participants made judgments about nine regions (harder task), individuals with more working memory capacity outperformed those with less working memory capacity. Other work finds that participants with low-working memory capacity make more accurate resource allocation decisions when using density plots and quantile dot plots compared to 95% confidence intervals, point estimates, or textual expressions of uncertainty [10]. Furthermore, participants with high-working-memory capacity were most accurate with quantile dot plots and reported less effort than all other tested methods. This work suggests that 95% confidence intervals, point estimates, are textual expressions of uncertainty require more working memory than densities and quantile dot plots [10]. The results of this study suggest that individual differences in working memory capacity primarily influence performance on complex decision-making tasks [10, 41].

4.5.2 Outlook on Using Cognitive Models in Visualization

Padilla et al.'s cognitive model [48] in Sect. 4.5.1 formalizes the implications of this dual mode of reasoning for visualization research. This cognitive model is an integration of multiple theories and takes a holistic approach to modeling decision-making with visualization. Applying this model can have a significant impact on design and evaluation of visualization interfaces. We provide some practical guidance for designers and visualization researchers on how to leverage visual features to generate Type 1 or Type 2 reasoning and evaluate visualization designs from a dual-process perspective.

One of the reasons why visualizations are so prominent is because they seem effortless. In other words, to design charts that bring about accurate, fast, and effortless reasoning, there needs to be a conscious effort to incorporate design principles that elicit bottom-up attention on task-relevant information. Padilla's model proposes that bottom-up attention is associated with Type 1 reasoning and top-down attention is more likely to generate Type 2 reasoning. Using this principle, Padilla et al. allow us to examine core design questions and provide guidelines to elicit either reasoning type by altering visual features.

Modeling visual attention is an important area of research in psychophysics, computational modeling, and neurophysiology (see a review of existing work by Borji and Itti [7]). When making a choice, the decision-maker must first decode the visualization via their visual system [70]. One way to elicit bottom-up attention is to align visual features to the users' existing graph schema. Figure 4.5 shows a figure from Padilla et al. where at first glance, it might appear that the introduction of the predator species caused a decline in the population of disease X [48]. If we look more closely at the graph, we notice that the y-axis is flipped and the predator species in fact contributed to the growth of species X. When decoding a visualization, we search our long-term memory for knowledge about how to interpret the chart and retrieve the graph schema that is the most similar. Altering graph conventions can cause errors because the graph schema will no longer match the chart. For example, multiple studies find that when the y-axis is inverted people

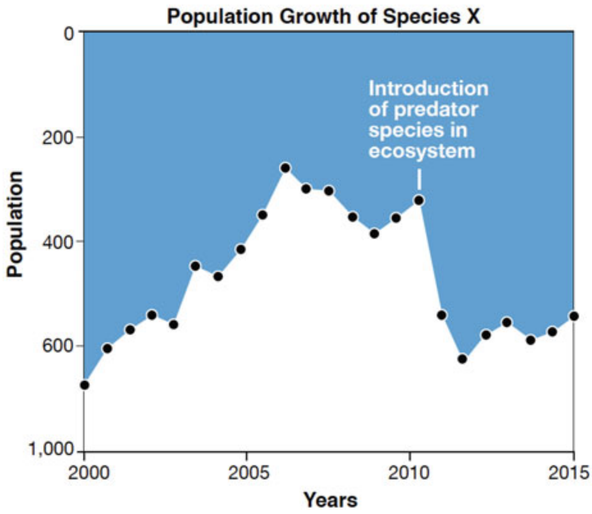

Fig. 4.5 Fictional relationship between the population growth of Species X and a predator species, where the Y-axis ordering does not match standard graphic conventions [48]

consistently come to the wrong interpretation of the chart [52, 72]. These errors are likely due to our reliance on graph schema to interpret graphs so much so that we do not notice when the schema does not match the chart.

One of the main design features that can affect decision type is saliency. Numerous studies showed that salient information in a visualization draws viewers' attention (e.g., [25, 25, 30, 45, 50, 61, 68]). First, it is important to identify the main piece of information that needs to be communicated and then we can direct the user's attention to this information using visual features. There exist behaviorally validated saliency models to determine the prominence of different visual encodings that will attract viewer's bottom-up attention, e.g., [28–30]. There is a long history of using saliency algorithms in computational imagery. For example, pioneering work by Koch and Ullmnan [39] created a *saliency map*—a two-dimensional topological map that encodes conspicuity across the entire scene. The central thesis of their work is that salient features within a stimulus "stand out," thus attracting overt attention. There have been some attempts in the visualization community to use this general principle to model visual attention in exploratory search tasks [45]. Still, future work is needed to model attention in the context of decision-making.

A critical component of Padilla et al.'s model is the principle that working memory is vital for Type 2 processing [48]. It is possible to gain insight into the type of decision-making generated by a visualization by measuring the user's working memory capacity. The amount of working memory generated by a task is commonly referred to as *cognitive load*. Remember that Type 1 reasoning does not require significant working memory contrarily to Type 2. There exists some prior work where the researchers have used measures of working memory to evaluate ease of use of visualization. Borgo et al. challenged traditional notions about chart junk and showed that embellishments do not generate higher cognitive load compared to other visualizations. By using a dual-task paradigm to evaluate different charts, they were able to evaluate differences in working memory elicited by different charts [6] by observing the dual-task cost. *Dual-task cost* is described as the decrease in performance between single and dual tasks. When the user completes two tasks simultaneously, significant memory is required, and by comparing dual-task cost across representations, differences in cognitive load can be inferred. There are a number of other ways to measure working memory. Castro et al. investigated the effect of various uncertainty visualizations on working memory using an operation span (OSPAN) task as part of a dual-task paradigm as well as self-reported measures [10]. They found that quantile dot plots and density plots are equally effective for low-working-memory individuals, while quantile dot plots elicit more accurate responses with less perceived effort for high-working-memory individuals. Moreover, Peck et al. used fNIRS to evaluate information visualization interfaces and found no difference in cognitive load in bar graphs and pie charts [54]. Other physical methods include *electroencephalogram (EEG)* [2] and *pupillometry*, which has shown high levels of correlation with working memory [47].

To summarize, two practical ways to elicit decision type are to design according to *graph schema* and *saliency*. For example, to elicit Type 1 reasoning, some elementary steps include verifying that your visualization does not violate any graphical conventions and brings forward important information using salient visual features. To examine decision type, one can observe *working memory* through self-reported measures, behavioral, and psychological methods. Padilla et al.'s model [48] is the most updated description of decision-making with visualizations, and we advocate that research incorporates this model when evaluating visualization design. Although we examined various decision-making models that appear in prior literature, they do not describe the entire visualization decision-making process using dual-process theory. For example, other models do not account for how framing effects of the visual or textual data might influence decisions [46]. Other factors such as individual differences (e.g., working memory capacity or spatial ability) can mediate the decision process [40, 44, 73] but are not encompassed in other models. Numerous researchers have voiced the importance of diversifying evaluation measures in the field of visualization [5], which is possible when using a cognitive framework. Ultimately, this chapter advocates for measures beyond the traditional usability measures, which capture *how* and *why* the brain processes visualizations.

4.6 Conclusion

Adopting decision models can have a significant impact on chart design and visualization evaluation. For instance, measuring working memory will diversify visualization research by tailoring chart design to individuals with varying levels of working memory capacity. Knowledge about dual-process reasoning and insight into cognitive load will enable tailoring visualization design to various tasks. We assert that for visualization to be reliably effective in real-world decision-making settings, research should consider leveraging existing decision theories when evaluating visual designs. We reviewed various utility-optimal theories, dual-process models, and cognitive science frameworks and discussed existing and future directions for visualization research. Much of the work discussed in this chapter raises valid concerns about evaluation paradigms that emphasize speed and accuracy measures. Overall, we advocate for evaluation techniques that go beyond traditional usability measures for better theoretical and practical advancements.

References

1. E. W. Anderson, K. C. Potter, L. E. Matzen, J. F. Shepherd, G. A. Preston, and C. T. Silva. A user study of visualization effectiveness using EEG and cognitive load. *Computer Graphics Forum*, 30(3):791–800, 2011.
2. E. W. Anderson, K. C. Potter, L. E. Matzen, J. F. Shepherd, G. A. Preston, and C. T. Silva. A user study of visualization effectiveness using EEG and cognitive load. *Computer Graphics Forum*, 30(3):791–800, 2011.
3. B. W. Balleine. The neural basis of choice and decision making. *Journal of Neuroscience*, 27(31):8159–8160, 2007.
4. M. Bancilhon, Z. Liu, and A. Ottley. Let's gamble: How a poor visualization can elicit risky behavior. *In proceedings of IEEE VIS Short Papers*, 2020.
5. E. Bertini, A. Perer, C. Plaisant, and G. Santucci. Beliv'08: Beyond time and errors: novel evaluation methods for information visualization. In *CHI'08 extended abstracts on Human factors in computing systems*, pages 3913–3916. 2008.
6. 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.
7. A. Borji and L. Itti. State-of-the-art in visual attention modeling. *IEEE transactions on pattern analysis and machine intelligence*, 35(1):185–207, 2012.
8. A. Bruhin, H. Fehr-Duda, and T. Epper. Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, 78(4):1375–1412, 2010.
9. S. Carpendale. Evaluating information visualizations. In *Information visualization*, pages 19–45. Springer, 2008.
10. S. Castro, P. S. Quinan, H. Hosseinpour and L. Padilla. (2022). Examining Effort in 1D Uncertainty Communication Using Individual Differences in Working Memory and NASA-TLX. *IEEE transactions on visualization and computer graphics, 28*(1), 411–421.
11. S. C. Castro, D. L. Strayer, D. Matzke, and A. Heathcote. Cognitive workload measurement and modeling under divided attention. *Journal of experimental psychology: human perception and performance*, 45(6):826, 2019.
12. W. S. Cleveland and R. McGill. An experiment in graphical perception. *International Journal of Man-Machine Studies*, 25(5):491–500, 1986.
13. W. S. Cleveland and R. McGill. Graphical perception: The visual decoding of quantitative information on graphical displays of data. *Journal of the Royal Statistical Society: Series A (General)*, 150(3):192–210, 1987.
14. N. Cowan. The many faces of working memory and short-term storage. *Psychonomic bulletin & review*, 24(4):1158–1170, 2017.
15. N. Cowan, J. S. Saults, and L. D. Nugent. The role of absolute and relative amounts of time in forgetting within immediate memory: The case of tone-pitch comparisons. *Psychonomic Bulletin & Review*, 4(3):393–397, 1997.

16. E. Dimara, S. Franconeri, C. Plaisant, A. Bezerianos, and P. Dragicevic. A task-based taxonomy of cognitive biases for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 26(2):1413–1432, 2020.
17. N. Elmqvist and J. S. Yi. Patterns for visualization evaluation. *Information Visualization*, 14(3):250–269, 2015.
18. J. S. B. Evans and K. E. Stanovich. Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, 8(3):223–241, 2013.
19. J. S. B. Evans and K. E. Stanovich. Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8(3):223–241, 2013.
20. D. Feldman-Stewart, M. D. Brundage, and V. Zotov. Further insight into the perception of quantitative information: Judgments of gist in treatment decisions. *Medical Decision Making*, 27(1):34–43, 2007.
21. M. Friedman and L. J. Savage. The utility analysis of choices involving risk. *Journal of political Economy*, 56(4):279–304, 1948.
22. M. Galesic and R. Garcia-Retamero. Using icon arrays to communicate medical risks: Overcoming low numeracy. *Health Psychology*, 28(2):210–216, 2009.
23. A. Hakone, L. Harrison, A. Ottley, N. Winters, C. Gutheil, P. K. Han, and R. Chang. Proact: iterative design of a patient-centered visualization for effective prostate cancer health risk communication. *IEEE transactions on visualization and computer graphics*, 23(1):601–610, 2016.
24. S. T. Hawley, B. Zikmund-Fisher, P. Ubel, A. Jancovic, T. Lucas, and A. Fagerlin. The impact of the format of graphical presentation on health-related knowledge and treatment choices. *Patient Education and Counseling*, 73:448–455, 2008.
25. M. Hegarty, A. Friedman, A. P. Boone, and T. J. Barrett. Where are you? the effect of uncertainty and its visual representation on location judgments in GPS-like displays. *Journal of experimental psychology: applied*, 22(4):381, 2016.
26. J. M. Hofman, D. G. Goldstein, and J. Hullman. How visualizing inferential uncertainty can mislead readers about treatment effects in scientific results. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2020.
27. W. Huang, P. Eades, and S.-H. Hong. Measuring effectiveness of graph visualizations: A cognitive load perspective. *Information Visualization*, 8(3):139–152, 2009.
28. L. Itti and C. Koch. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision research*, 40(10–12):1489–1506, 2000.
29. L. Itti and C. Koch. Computational modelling of visual attention. *Nature reviews neuroscience*, 2(3):194–203, 2001.
30. L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence*, 20(11):1254–1259, 1998.
31. S. L. Joslyn and J. E. LeClerc. Uncertainty Forecasts Improve Weather-Related Decisions and Attenuate the Effects of Forecast Error. *Journal of Experimental Psychology: Applied*, 18(1):126–140, 2012.
32. D. Kahneman. A perspective on judgment and choice: mapping bounded rationality. *American psychologist*, 58(9):697, 2003.
33. D. Kahneman. *Thinking Fast and Slow*. Farrar, Straus and Giroux, New York, 2011.
34. D. Kahneman and S. Frederick. Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment*, 49:81, 2002.
35. D. Kahneman and A. Tversky (2013) Prospect Theory: An Analysis of Decision under Risk. In: MacLean, L.C. and Ziemba, W.T., Eds., Handbook of the Fundamentals of Financial Decision Making: Part I, World Scientific Publishing, Hackensack, 99–127.
36. A. Kale, M. Kay, and J. Hullman. Visual reasoning strategies for effect size judgments and decisions. *IEEE Transactions on Visualization and Computer Graphics*, 2020.
37. M. Kay, T. Kola, J. Hullman, and S. Munson. When (ish) is my bus?: User-centered visualizations of uncertainty in everyday mobile predictive systems. *In Proceedings of the CHI Conference on Human Factors in Computing System*, page 5092–5103, 2016.

38. R. L. Keeney and H. Raiffa. *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge university press, 1993.
39. C. Koch and S. Ullman. Shifts in selective visual attention: towards the underlying neural circuitry. In *Matters of intelligence*, pages 115–141. Springer, 1987.
40. Z. Liu, R. J. Crouser, and A. Ottley. Survey on individual differences in visualization. *Computer Graphics Forum*, 39(3):693–712, 2020.
41. G. L. Lohse. The role of working memory on graphical information processing. *Behaviour & Information Technology*, 16(6):297–308, 1997.
42. G. A. Miller. *11. The magical number seven, plus-or-minus two or some limits on our capacity for processing information*. University of California Press, 2020.
43. T. Munzner. A nested model for visualization design and validation. *IEEE transactions on visualization and computer graphics*, 15(6):921–928, 2009.
44. A. Ottley. Adaptive and personalized visualization. *Synthesis Lectures on Visualization*, 7(1):1–117, 2020.
45. A. Ottley, R. Garnett, and R. Wan. Follow the clicks: Learning and anticipating mouse interactions during exploratory data analysis. *Computer Graphics Forum*, 38(3):41–52, 2019.
46. A. Ottley, E. M. Peck, L. T. Harrison, D. Afergan, C. Ziemkiewicz, and R. Taylor, H. A.... & Chang. Improving Bayesian reasoning: The effects of phrasing, visualization, and spatial ability. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):529–538, 2015.
47. L. M. Padilla, S. C. Castro, P. S. Quinan, I. T. Ruginski, and S. H. Creem-Regehr. Toward objective evaluation of working memory in visualizations: a case study using pupillometry and a dual-task paradigm. *IEEE transactions on visualization and computer graphics*, 26(1):332–342, 2019.
48. L. M. Padilla, S. H. Creem-Regehr, M. Hegarty, and J. K. Stefanucci. Decision making with visualizations: A cognitive framework across disciplines. *Cognitive Research: Principles and Implications*, 3(1):1–25, 2018.
49. L. M. Padilla, M. Powell, M. Kay, and J. Hullman. Uncertain about uncertainty: How qualitative expressions of forecaster confidence impact decision-making with uncertainty visualizations. *Frontiers in Psychology*, 11, 2020.
50. L. M. Padilla, I. T. Ruginski, and S. H. Creem-Regehr. Effects of ensemble and summary displays on interpretations of geospatial uncertainty data. *Cognitive research: principles and implications*, 2(1):1–16, 2017.
51. L. M. K. Padilla, S. C. Castro, P. S. Quinan, I. T. Ruginski, and S. H. Creem-Regehr. Toward objective evaluation of working memory in visualizations: A case study using pupillometry and a dual-task paradigm. *in IEEE Transactions on Visualization and Computer Graphics*, 26(1):332–342, Jan. 2020.
52. A. V. Pandey, K. Rall, M. L. Satterthwaite, O. Nov, and E. Bertini. How deceptive are deceptive visualizations? an empirical analysis of common distortion techniques. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1469–1478, 2015.
53. E. M. Peck, B. F. Yuksel, L. Harrison, A. Ottley, and R. Chang. Position paper: towards a 3-dimensional model of individual cognitive differences. *Beyond Time & Errors–Novel Evaluation Methods for Visualization (Proceedings of BELIV 2012)*, 2012.
54. E. M. M. Peck, B. F. Yuksel, A. Ottley, R. J. Jacob, and R. Chang. Using fNIRS brain sensing to evaluate information visualization interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 473–482. ACM, 2013.
55. S. Pinker. A theory of graph comprehension. *Artificial intelligence and the future of testing*, pages 73–126, 1990.
56. C. Plaisant. The challenge of information visualization evaluation. In *Proceedings of the working conference on Advanced visual interfaces*, pages 109–116, 2004.
57. G. Loewenstein and D. Prelec. (1991). Decision Making Over Time and Under Uncertainty: A Common Approach. Management Science. 37. 770–786. 10.1287/mnsc.37.7.770.
58. V. Reyna and C. Brainerd. Fuzzy-trace theory: An interim synthesis. *Learning and individual Differences*, 7(1):1–75, 1995.

59. V. Reyna and F. Lloyd. Physician decision making and cardiac risk: Effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *Journal of Experimental Psychology: Applied*, 12(3):179–195, 2006.
60. V. F. Reyna. A theory of medical decision making and health: fuzzy trace theory. *Medical Decision Making*, 28(6):850–865, 2008.
61. I. T. Ruginski, A. P. Boone, L. M. Padilla, L. Liu, N. Heydari, H. S. Kramer, M. Hegarty, W. B. Thompson, D. H. House, and S. H. Creem-Regehr. Non-expert interpretations of hurricane forecast uncertainty visualizations. *Spatial Cognition & Computation*, 16(2):154–172, 2016.
62. J. G. Ruiz, A. D. Andrade, R. Garcia-Retamero, R. Anam, R. Rodriguez and J. Sharit. (2013). Communicating global cardiovascular risk: are icon arrays better than numerical estimates in improving understanding, recall and perception of risk?. *Patient education and counseling, 93*(3), 394–402.
63. T. L. Saaty and M. S. Ozdemir. Why the magic number seven plus or minus two. *Mathematical and computer modelling*, 38(3–4):233–244, 2003.
64. B. Saket, A. Endert, and J. Stasko. Beyond usability and performance: A review of user experience-focused evaluations in visualization. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization*, pages 133–142, 2016.
65. D. Simkin and R. Hastie. An information-processing analysis of graph perception. *Journal of the American Statistical Association*, 82(398):454–465, 1987.
66. I. Spence and S. Lewandowsky. Displaying proportions and percentages. *Applied Cognitive Psychology*, 5(1):61–77, 1991.
67. K. E. Stanovich. *Who is rational?: Studies of individual differences in reasoning*. Psychology Press, 1999.
68. E. R. Stone, W. R. Sieck, B. E. Bull, J. F. Yates, S. C. Parks, and C. J. Rush. Foreground: background salience: Explaining the effects of graphical displays on risk avoidance. *Organizational behavior and human decision processes*, 90(1):19–36, 2003.
69. A. Tversky. A critique of expected utility theory: Descriptive and normative considerations. *Erkenntnis*, 9(2):163–173, 1975.
70. C. Ware. *Visual thinking for design*. Elsevier, 2010.
71. S. J. Westwood, S. Messing, and Y. Lelkes. Projecting confidence: How the probabilistic horse race confuses and demobilizes the public. *The Journal of Politics*, 82(4):1530–1544, 2020.
72. G. Woodin, B. Winter and L. Padilla, "Conceptual Metaphor and Graphical Convention Influence the Interpretation of Line Graphs," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 2, pp. 1209–1221, 1 Feb. 2022.
73. C. Ziemkiewicz, A. Ottley, R. J. Crouser, K. Chauncey, S. L. Su, and R. Chang. Understanding visualization by understanding individual users. *IEEE computer graphics and applications*, 32(6):88–94, 2012.