The Anatomy of a Plea: How Uncertainty, Visualizations & Individual Differences Shape Plea Bargain Decisions

Melanie Bancilhon
m.bancilhon@gwu.edu
US DEVCOM Army Research Laboratory
Washington, DC, USA

Alvitta Ottley
alvitta@wustl.edu
Washington University in St. Louis
St. Louis, MO, USA

Andrew Jordan
awjordan@wustl.edu
Washington University in St. Louis
St. Louis, MO, USA

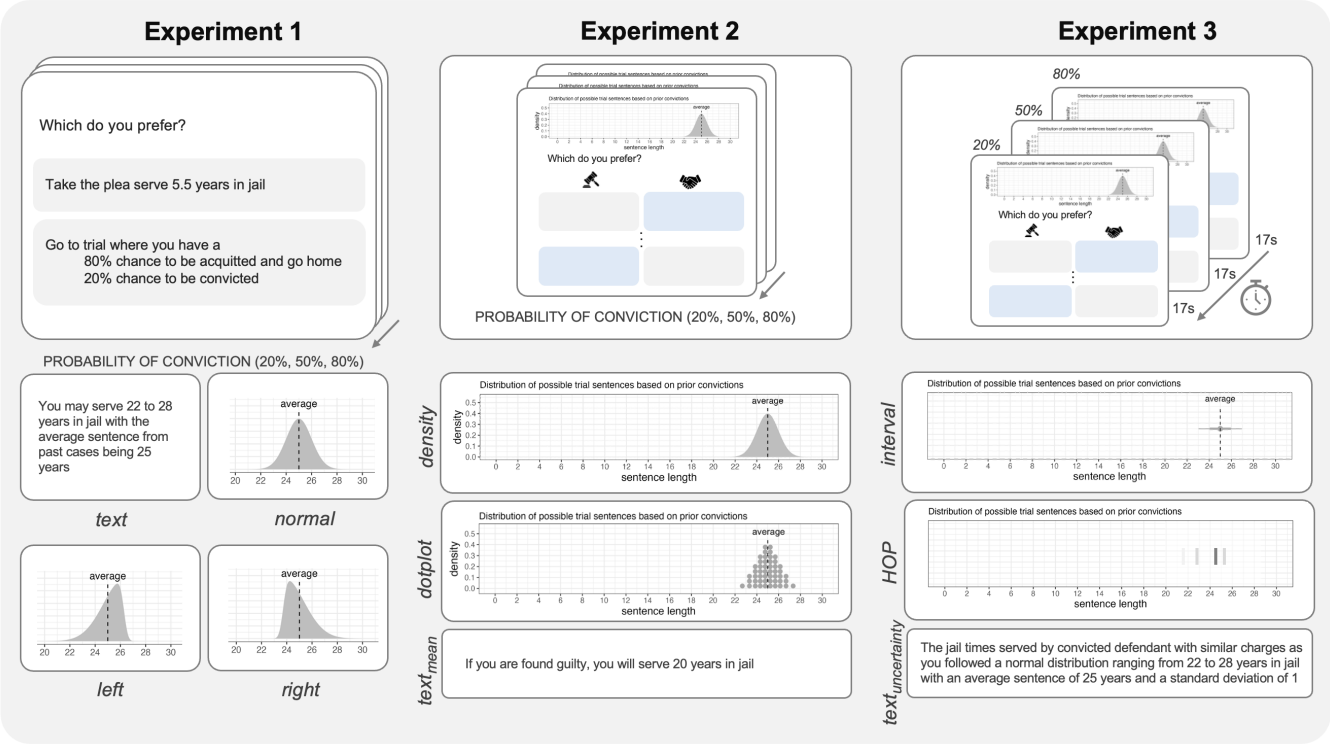

Figure 1: 1) Experiment 1 consisted of a series of 12 choices where participants were asked to decide between taking the plea or going to trial. Each participant saw distributions of potential trial sentences based on prior convictions via four presentation conditions: *text*, *normal*, *left*, and *right*. Each format conveyed 3 probabilities of conviction at trial: 20%, 50%, and 80% for a total of 12 trials. 2) In Experiment 2, participants were shown the distribution of potential trial sentences based on prior convictions via one of six randomly assigned presentation conditions: *density*, *dotplot*, *interval*, *HOP*, $text_{uncertainty}$ or $text_{mean}$. They were presented with a decision sheet where they could choose when to take the plea and when to switch their decision to enter the trial. The study consisted of 3 decision sheets corresponding to the probability of conviction at trial values of 20%, 50%, and 80%. 3) Experiment 3 was identical to Experiment 2, but participants only had 17 seconds to enter their choices.

CHI '25, April 26-May 1, 2025, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713096

ABSTRACT

Plea bargains are commonly used in the criminal justice system, where they can offer potential benefits to both the prosecution and the defendant. However, research has shown that defendants often engage in poor decision-making, such as accepting the plea even when the trial sentence is likely to be less severe. While previous studies have shown some evidence that uncertainty visualizations can improve decision-making, there is a lack of research on their

effectiveness in domain-specific tasks like plea bargain decision-making. In this work, we conduct a series of experiments to explore whether the presence and format of uncertainty impact plea bargain decisions, taking into account time pressure and individual differences. Our findings reveal that these factors can have a significant impact on plea bargain decisions. We also show evidence that communicating uncertainty in the form of text can elicit more optimal decisions under time-pressure conditions.

CCS CONCEPTS

• **Human-centered computing → Information visualization**;
• **Applied computing → Law**.

KEYWORDS

Plea Bargain, Uncertainty Visualizations, Decision-Making, Criminal Justice, Individual Differences, Time Pressure

ACM Reference Format:
Melanie Bancilhon, Alvitta Ottley, and Andrew Jordan. 2025. The Anatomy of a Plea: How Uncertainty, Visualizations & Individual Differences Shape Plea Bargain Decisions. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3706598.3713096

1 INTRODUCTION

In the US, 75% to 90% of criminal cases are solved through plea bargaining, which accounts for the vast majority of guilty sentences [1, 71]. This agreement between the prosecution and the defendant may allow defendants to avoid the risk of conviction at trial on a more serious charge. Plea deals are typically offered with a time constraint, after which a worse deal is offered. The "shadow of trial" theory suggests that plea bargain decisions are made based on the expected trial outcome, whereby if the expected penalty after a trial is significantly harsher than the plea offer, defendants should accept the plea.

However, in reality, there exists many circumstantial factors and cognitive biases that can affect plea bargain decision-making, often leading to sub-optimal choices [16, 28, 30, 46, 66]. For example, some defendants may accept plea bargains even if they would not have been charged or given such harsh sentencing if they had gone to trial [4, 9]. While prior work suggests that effectively communicating uncertainty can improve decision-making, there is a lack of empirical evaluation of the effects of uncertainty communication and visualization on decision-making tasks in real-life applications.

We present three empirical experiments demonstrating that skewness, uncertainty communication, and visualization design can impact plea bargain decisions, and that these effects are shaped by factors such as time pressure and individual differences. Our findings revealed that participants adjusted their decisions based on the skewness of the distribution of trial sentences, making more risk-averse decisions with the left-skewed density chart than the chart depicting a normal distribution. We also demonstrated that numeracy and demographic factors impacted plea bargain decisions, and general risk propensity impacted decision confidence. Finally, when subject to time pressure, text communicating uncertainty elicited more risk-optimal decisions than text communicating only the mean and uncertainty visualizations. This provides evidence that text can sometimes be better than visualization for decision-making.

Altogether, our work has theoretical implications for visualization research and practical implications for developing plea bargain decision aids. Based on the results of this work, we encourage the visualization community not to overlook text as a viable format for decision support and to account for the influence of situational factors and individual differences when assessing the effectiveness of visualizations, especially across domain-specific applications.

SUMMARY OF FINDINGS

Across experiments 1, 2, and 3, the likelihood of being guilty at trial significantly impacted plea bargain decisions. Participants made more risk-averse decisions when the probability of conviction at trial was higher.

Experiment 1 showed that skewness has a significant impact on plea bargain decisions, where the left-skewed chart elicited significantly more trial decisions, i.e., risk-optimal decisions, compared to the normal chart.

Experiment 2 showed no significant impact of the presence and format of uncertainty on plea bargain decisions. However, participants with high numeracy made more risk-neutral decisions, and those with high general risk propensity (GRP) reported higher confidence in their decisions. Numeracy and GRP were negatively correlated. We found no significant relationship between demographic factors or precedents and decision-making.

Experiment 3 is identical to Experiment 2 but participants had an added time constraint to make their decisions. Our findings revealed that participants made significantly less optimal decisions overall when subject to time pressure. Moreover, we found that under time pressure conditions, uncertainty text elicited significantly more risk-optimal decisions compared to mean-only text as well as all uncertainty visualizations. The effect of numeracy and GRP on decision and decision confidence was consistent with Experiment 2. Follow-up analyses with interaction effects revealed that when using Hypothetical Outcome Plots (*HOPs*), participants with higher numeracy scores made more risk-optimal decisions compared to those with low numeracy scores. We also found that demographic factors were associated with decisions, where older participants and Black/African American participants tended to make less risk-optimal decisions.

2 RELATED WORK

A plea bargain is a negotiated agreement between the prosecutor and a defendant to resolve a criminal case without going to trial. The defendant typically agrees to plead guilty to a lesser charge or receive a more lenient sentence. According to the American Bar Association, plea bargains account for about 98% of federal convictions and 95% of state convictions in the United States [5]. Supporters of plea bargains argue that they speed up the legal process and alleviate the strain on the U.S. criminal justice system [5]. Despite these benefits, there is substantial evidence that defendants frequently

make poor decisions due to inadequate legal representation, lack of access to information, lack of support for decision-making, and situational factors such as stress, time constraints, and personal circumstances [5]. This work aims to contribute to the design of decision aids to support defendants as they consider sentencing outcomes when choosing between maintaining their right to a trial or accepting a plea deal.

2.1 Factors That Influence Plea Bargain Decisions

Circumstantial Factors & Cognitive Biases. Many factors that can affect plea bargain decision-making. Prior work has shown that the probability of conviction at trial can impact plea bargain decisions [56, 66]. For example, Small et al. conducted a study where they used probability discounting to study the choice between accepting a plea bargain or risking a trial. They found that maximum potential sentence length did not impact plea bargain acceptance, but the likelihood of conviction did [66]. Other circumstantial factors, such as whether the defendant is innocent or guilty, can also impact how they decide about plea bargains [16, 26, 30]. Prior work also found that cognitive biases such as risk aversion [21], temporal discounting [53], and cognitive processing style [30] can affect plea bargain decisions. Helm et al. found that the type of mental representation that an individual relies on, whether gist or verbatim, impacts plea bargain decisions [30]. Fuzzy Trace Theory, which posits that people rely on two types of memory representations—precise, detailed "verbatim" memories and general, meaning-based "gist" memories—provides a framework for understanding these findings [58]. For example, Helm et al. found that individuals relying more on gist are influenced by meaning-based distinction, such as being guilty versus innocent or being charged with a felony versus a misdemeanor. On the other hand, individuals relying on verbatim representations engage in a rational reasoning process that does not reflect their underlying values [30].

Age & underdeveloped cognitive faculties. Demographic factors such as age can also impact plea bargain decisions. Gollwitzer et al. found that younger adults were more likely to make decisions based on whether they were guilty or innocent – they would take the plea if they were guilty and go to trial if they were innocent [27]. This could be due to an increased propensity for impulsivity, poor decision-making [67], heightened just-world beliefs [61], or amplified illusion of transparency [14] during emerging adulthood.

Time pressure and stress. Plea bargains are often subject to a time constraint, after which a worse deal is typically offered [5]. Investigations into the effect of time pressure on decision-making highlight divergent perspectives, where time-dependent incentives sometimes lead to worse performance [55] and sometimes lead to faster and better decisions [41]. Some researchers have examined the effect of uncertainty and visualization on judgment under time pressure [20, 42, 59]. For example, using a hypothetical helicopter rescue scenario, Korporaal et al. found that neither time pressure nor uncertainty affected participants' decision-making accuracy. However, uncertainty affected their decision strategies, and time pressure affected their response times [42]. Cheong et al. investigated the effect of different uncertainty maps on emergency decision-making and found that while visualization did not affect performance under normal conditions, boundary visualizations were worse under time pressure [20].

Taken together, considerable evidence suggests that plea bargain decision-making and situational circumstances are often less than ideal and can lead to injustices. This work primarily addresses issues related to data access and misunderstandings. In particular, it aims to help defendants make more informed decisions and better consider the consequences associated with rejecting a plea bargain. Exploring how data visualization can convey sentencing information can result in better access and understanding of data, which could contribute to more informed and better decisions.

2.2 Data Uncertainty, Visualization, & Decision-Making

Sentencing data inherently contains uncertainty, and many researchers argue it is important to use transparent visual representations to convey uncertainty accurately. This has led to a large body of work focused on designing and comparing the effectiveness of various representations of uncertainty. Prior work has shown that design choices for uncertainty visualizations can impact accuracy [24, 32, 33, 39, 40, 52, 62] and recall [32]. For example, uncertainty visualizations that encode information using natural frequencies, such as *quantile dotplots*, have led to better performance than other charts across several studies.

Hullman et al. demonstrated that graphical uncertainty prediction techniques and discrete visualizations, such as quantile dotplots and interval plots, significantly improved users' recall and estimation of effect uncertainty compared to traditional methods like bar charts and density plots [32]. Quantile dotplots were shown to reduce the variance of probabilistic estimates, making them a robust tool for visualizing uncertainty in continuous data distributions [40]. Hypothetical outcome plots (HOPs), which display a series of possible outcomes from a statistical distribution animated over time, have been shown to enable more accurate inferences about uncertain information to static representations [33, 39].

However, whether communicating data uncertainty results in better decisions than point estimates is disputed, and very few works explicitly compare the two. Some studies have shown that communicating uncertainty instead of single point estimates can improve reasoning accuracy and decision-making [22, 35, 36, 49, 54, 60]. For example, Joslyn and LeClerc found that displaying uncertainty in weather predictions can lead to more optimal decision-making and trust in forecasts [35]. In contrast, prior work has shown evidence that understanding uncertainty is challenging for novices and experts alike [11, 17, 31, 37, 65]. Several studies that examine data workers' perception and use of uncertainty highlight the difficulty of balancing transparency, accuracy, and simplicity [17, 31, 37, 65].

Despite the existing body of work, there are important unanswered questions regarding the potential effectiveness of uncertainty and visualizations in the context of plea bargain decisions. Specifically, we have limited knowledge on whether communicating uncertainty leads to more optimal decision-making. Moreover,

only a few studies have examined the impact of uncertainty visualizations on decision-making [22, 35, 37]. In their work, Kale et al. found that visualizations that support the least biased effect size estimation do not support the best decision-making [37], highlighting the importance of expanding current evaluation methods to decision tasks. Finally, we lack knowledge on how uncertainty and visualizations can impact plea bargain decisions. These questions are the primary focus of this paper. In addition, we will examine the impact of time pressure, which is a crucial characteristic of plea bargains, as prosecutors typically impose time limits on their offers.

3 GOALS, PRELIMINARY DATA ANALYSIS & PRE-REGISTRATION

This project is part of a larger effort to develop visual representations of sentencing trends and models to assist defendants in making plea bargain decisions and is inspired by our conversations with public defenders, prison wardens, and current inmates. To explore how to represent sentencing data, we collected and examined data from Cook County, Illinois, covering the period from 1984 to 2016. We selected Cook County because it encompasses the Chicago metropolitan area, which is a high-crime region in the US and representative of a considerable population of people who make plea bargain decisions. This dataset included information about 384,917 defendants, including their demographics, charges, and sentencing details. Our initial analysis was conducted to better understand the real-world characteristics of the data in order to ensure the validity of our research questions and methodology. We specifically aimed to grasp the distributions of sentencing data, the prevalence of crimes, and the demographics of the defendants.

Unlike previous studies focusing on the impact of uncertainty visualization on accuracy and decisions, the plea bargain scenario is unique in that it involves a combination of factors to consider. These aspects include circumstantial factors such as the likelihood of conviction at trial or defendant innocence [16, 26, 30, 56, 66], various cognitive biases [21, 30, 53], individual characteristics of defendants (particularly emerging adults and individuals with a history of trauma) [14, 27, 61, 67], and the time-sensitive nature of these decisions [41, 55]. Furthermore, the existing literature does not provide enough evidence on how and whether data distributions affect decisions, the influence of visual representation on individual characteristics in plea bargain decisions, and the impact of time pressure on decision-making with data visualizations.

To address this knowledge gap, we conducted three pre-registered experiments[1,2,3] to investigate how visual representations of sentencing data might affect the way people make plea bargain decisions. Although our pre-registered analysis plan proposed using non-parametric tests for all three experiments, we opted to conduct regression models (logistic and linear mixed effect models) instead. This decision was driven by the hierarchical structure of the data, with repeated measures nested within participants, where regressions are better suited to the model. Additionally, regressions include random effects, offering a more robust framework for hypothesis testing [48]. Before the regressions, we applied Box-Cox power transformations where applicable to approximate a normal distribution, a step that was also not included in our pre-registrations. All data and analyses are included in our osf project.

- *Experiment 1.* Prior work typically assumes a normal data distribution when examining the impact of visualization on decisions. Sentencing data often deviate from a normal distribution because of minimum and maximum sentencing requirements. We relax this assumption and examine how the skewness of the underlying data might affect plea bargain decisions with uncertainty visualization.[1]
- *Experiment 2.* We compare plea bargain decisions using four uncertainty visualizations and two text-only representations. Also, we examine how numerical skills and general risk propensity impact these decisions and how demographics might be associated with plea bargain decisions. The ethnicity and age of the individuals in our study were chosen to reflect the makeup of the Cook County (Chicago Metropolitan region) defendant population.[2]
- *Experiment 3.* We consider the impact of time pressure on decisions.[3]

4 EXPERIMENT 1: THE EFFECT OF SKEWNESS

The existing research on the impact of uncertainty visualization on decisions assumes that the underlying data is Gaussian with moderate variance. However, the distribution of trial sentences does not typically follow a normal distribution and is influenced by policies such as minimum and maximum sentencing. Our analysis of data from the Chicago Cook County court revealed a right-skewed distribution, where the majority of defendants received shorter sentences, while only a few were given longer trial sentences. Furthermore, research in behavioral economics theorizes that the findings of previous studies on how visualizations of normally distributed data affect decisions may not be applicable when the data is skewed [10]. One study showed that positively skewed distributions were preferred in a lottery game because people were attracted to large, although unlikely gains [70]. Conversely, negatively skewed distributions were avoided due to the fear of significant loss [70]. Thus, **Experiment 1 examines how the skewness of the distribution of potential trial sentences upon conviction biases plea bargain decisions.**

4.1 Stimuli

We evaluate a *normal* distribution, *left*-skewed distribution, and *right*-skewed distribution depicted using a probability density graph (see figure 1). We used *text* as a control condition. All conditions had a standard deviation of 1 for the corresponding mean (see figure 2). The *normal* condition was Gaussian with moderate variance, while the skewed conditions had underlying skewed-normal functions with a skewness factor of 0.9 for *right* and -0.9 for *left*. In each of the graphs, the mean was explicitly indicated. The x-axis was labeled "sentence length", and the y-axis was labeled "density" as exampled in Figure 1 (see supplementary material for stimuli).

[1] Experiment 1 pre-registration
[2] Experiment 2 pre-registration
[3] Experiment 3 pre-registration

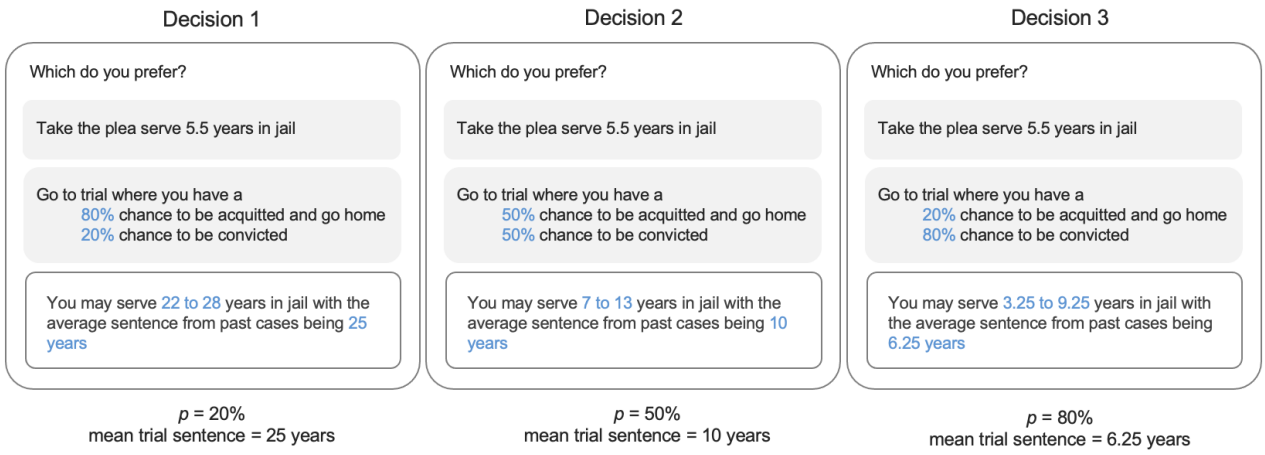

Figure 2: Set-up for Experiment 1 showing 20%, 50% and 80% probability of conviction at trial using *text*. The plea and trial sentences were determined such that the expected value of the trial was always 0.5 years less than the plea, making the former the most optimal choice across all rounds. In our study, participants saw 4 skewness conditions, i.e., *text*, *normal*, *left*, and *right* for a total of 12 rounds.

Category	Demographics	Exp 1	Exp 2	Exp 3
Gender	Male	55.59%	88.25%	88.83%
	Female	41.93%	10.89%	10.53%
	Non-Binary	2.17%	0.57%	0.64%
Race	Black	18.32%	75.58%	74.68%
	White	64.91%	13.37%	13.72%
	Asian	6.83%	–	–
	Hispanic	–	10.17%	10.64%
	Mixed	7.76%	–	–
	Other	2.80%	0.87%	0.96%
Age	18-34	52.48%	36.39%	47.13%
	35-54	39.13%	52.72%	46.17%
	55+	8.39%	10.89%	6.70%
Education	< High School	1.24%	1.43%	0.53%
	High School	36.96%	32.09%	32.02%
	Bachelor's	40.37%	42.41%	49.68%
	Master's	17.70%	20.92%	15.32%
	Doctoral	3.73%	3.15%	2.45%

Table 1: Demographics for Experiments 1, 2, and 3 as collected by Prolific. In Experiments 2 and 3, we recruited a custom sample from Prolific (based on gender and ethnicity) to better reflect the defendant population as recorded by the Cook County Court [34]. Note that the race categories for Experiment 1 differ from Experiments 2 and 3 since they include the default categories defined by Prolific, i.e., White, Black, Asian, Mixed, and Other.

4.2 Participants

We recruited 323 participants from Prolific. After excluding participants who took less than 3 seconds on average to make decisions, we were left with 322 participants. Each user saw all three visualization conditions (*normal*, *left* and *right*) as well as the *text* condition, making our manipulations of the effect of visualization all within-subject. Users completed randomly ordered 12 trials, and an attention check trial was conducted in the middle of the block of trials. In the attention check trial, users saw a probability of conviction of 100%, making the plea deal the obvious best choice. We did not observe a learning effect in this study (see supplementary material for analysis). We estimated the task to last around 20 minutes. Participants received a base pay of $8 and a bonus that depended on their performance on the task. They received $12 initially, and for each year of jail time received, $0.10 was subtracted from that amount. The median time taken across all participants was 20.18 minutes with a standard deviation of 14.14. The median bonus across all participants was $7.20, with a standard deviation of 1.38.

4.3 Task & Procedures

After agreeing to participate and reviewing the instructions, participants were presented with a hypothetical scenario. They were asked to imagine they were facing drug charges and had been offered a plea deal. Participants then completed 12 experimental questions and 1 attention question for a total of 13 decision questions. For each decision, participants were provided with information about a plea deal and the potential outcomes of choosing to go to trial. They had to choose between two options: (1) Plea Deal: Accept a predetermined legal outcome, or (2) Trial: Risk an uncertain outcome by going to trial. This study was a within-subject experiment design where for each skewness condition (*text*, *normal*, *left* and *right*), participants completed three decision questions with 20%, 50% and

80% probability of conviction at trial (see figure 2). The average trial sentences were set such that the expected value of the trial sentence was 5 years for all decision questions. The plea bargain sentence was held constant at 5.5 years. Our choice for the plea sentence and the expected value of the trial sentence, whereby the trial is the risk optimal decision, was informed by theoretical and empirical evidence showing that defendants are often offered plea deals that are greater than the expected trial sentence. [2, 43]. A risk-neutral agent values a trial at the mean of the sentencing distribution and will, therefore, choose the trial if the expected sentence is less than the plea. Therefore in this scenario, the utility-optimal decision is always the trial. Throughout Experiment 1, we refer to trial decisions as *risk optimal* decisions.

If participants chose the trial, they were communicated the trial outcome, sentence length, and remaining bonus, simulated using the corresponding trial data. In this simulation, conducted each time the user selects the trial, the trial outcome is determined by probabilistically assigning the defendant an innocent or guilty verdict. The probability of a guilty verdict corresponds to the probability of conviction for the current decision. In the case of an innocent verdict, the sentence length is set to 0. In the case of a guilty verdict, the sentence length is sampled from the probabilistic distribution of sentence lengths for the current decision. Probabilistic distributions are skewed for the skewed conditions and Gaussian for the text and normal conditions. If a participant chooses the plea, the plea sentence and remaining bonus will be communicated. For every year in jail, $0.10 was deducted from their initial bonus. After each plea bargain decision, we asked participants to estimate on Likert scales their perceived likelihood of conviction *("What was your likelihood of conviction at trial?")*, decision confidence *("How confident are you in your decision")*, perceived trial sentence severity *("Rate the severity of the possible trial sentence")*, perceived plea sentence severity *("Rate the severity of the possible plea deal sentence")* and justification for their decision *("Why did you accept or reject the plea bargain?")*. After completing the trials, participants were asked to answer some demographic questions about their gender, ethnicity, age, and education level. They were also asked whether they had been accused or convicted of a drug crime and if they were ever offered a plea bargain.

4.4 Analysis

We conducted a mixed-effect logistic model to investigate whether skewness and probability influence plea bargain decision (*decision*) and a linear mixed effect model (LME) to examine whether they impact the perception of the likelihood of conviction (*perceptionOfLikelihood*), perception of severity ratio (*severityRatio*) and decision confidence (*decisionConfidence*). The models included p and *format* as fixed effects and participant identifier *id* as a random intercept to account for repeated measures within subjects. The reference level for the models was *normal*. The outcome variables were transformed using *boxcox* power transformations to improve normality. We report the significant findings from our analysis. We also conducted mixed-effects logistic models with interaction effects for follow-up analyses.

- $Exp1_{decision}$: $decision \sim \text{glmer}(p + format + (1 \mid id))$
- $Exp1_{confidence}$: $decisionConfidence \sim \text{lmer}(p + format + (1 \mid id))$
- $Exp1_{likelihood}$: $perceptionOfLikelihood \sim \text{lmer}(p + format + (1 \mid id))$
- $Exp1_{severity}$: $severityRatio \sim \text{lmer}(p + format + (1 \mid id))$

where:

$decision$	$\in \{0, 1\}$; choice between plea and trial
$decisionConfidence$	$\in [0...10]$; confidence in the decision
$perceptionOfLikelihood$	$\in [0...10]$; perceived likelihood of going to trial
$severityRatio$	$\in [0...1]$; ratio of perceived severity of trial to plea

4.5 Results

Probability of Conviction. The proportion of trial decisions, i.e., risk optimal decisions across probability of conviction 20%, 50%, and 80% were 71.3%, 52.9%, and 61.6%, respectively. Model $Exp1_{decision}$ revealed a significant effect of the probability of conviction on plea bargain decisions where 50% (Estimate = -0.991, SE = 0.093, z = -10.641, $p < 0.001$) and 80% (Estimate = -0.543, SE = 0.093, z = -5.827, $p < 0.001$) were significantly associated with a decrease in the likelihood of a trial decision compared to the baseline of 20%. Pairwise comparisons revealed that across all four presentation conditions, differences between all three values were significantly different with the exception of 50% and 80% in the *left* condition.

We found a significant effect of probability of conviction on perceived likelihood of conviction, where 50% (Estimate = 1.53, SE = 0.034, z = 44.61, $p < 0.001$) and 80% (Estimate = 2.82, SE = 0.034, z = 82.36, $p < 0.001$) elicited significantly higher ratings of perceived likelihood of conviction at trial compared to 20%. We also found a significant effect of probability of conviction on severity ratio where participants rated the trial sentence to be more severe than the plea sentence for both 50% (Estimate = -0.3222, SE = 0.0135, z = -23.92, $p < 0.001$) and 80% (Estimate = -0.4764, SE = 0.0135, z = -35.36, $p < 0.001$). Finally, we found a significant effect of probability of conviction on decision confidence where compared to the 20% baseline, participants reported significantly less confidence in their decisions at 50% (Estimate = -1.35, SE = 0.01837, z = -7.36, $p <$ 0.001) and significantly more confidence in their decisions at 80% (Estimate =0.43, SE = 0.1837, z = 2.35, $p < 0.05$).

Our results show that participants are more likely to choose the plea bargain when the probability of conviction at trial is large, despite the trial having a smaller expected sentence. However, participants made fewer trial decisions when the probability of conviction was 50% compared to 80%.

Skewness. The proportion of trial decisions, i.e., risk optimal decisions across *left*, *right*, *normal*, and *text* conditions were 58.7%, 65.2%, 62.8%, and 60.9% respectively. Model $Exp1_{decision}$ revealed that the *left* condition elicited significantly fewer trial (risk optimal) decisions compared to the baseline condition *normal*. We found no significant impact of skewness on the perceived likelihood of conviction, relative severity ratio, or decision confidence. *Our results show that despite the expected trial sentence consistently being less than the plea sentence, participants are more likely to choose the trial when its sentencing distribution is left-skewed compared to normal.*

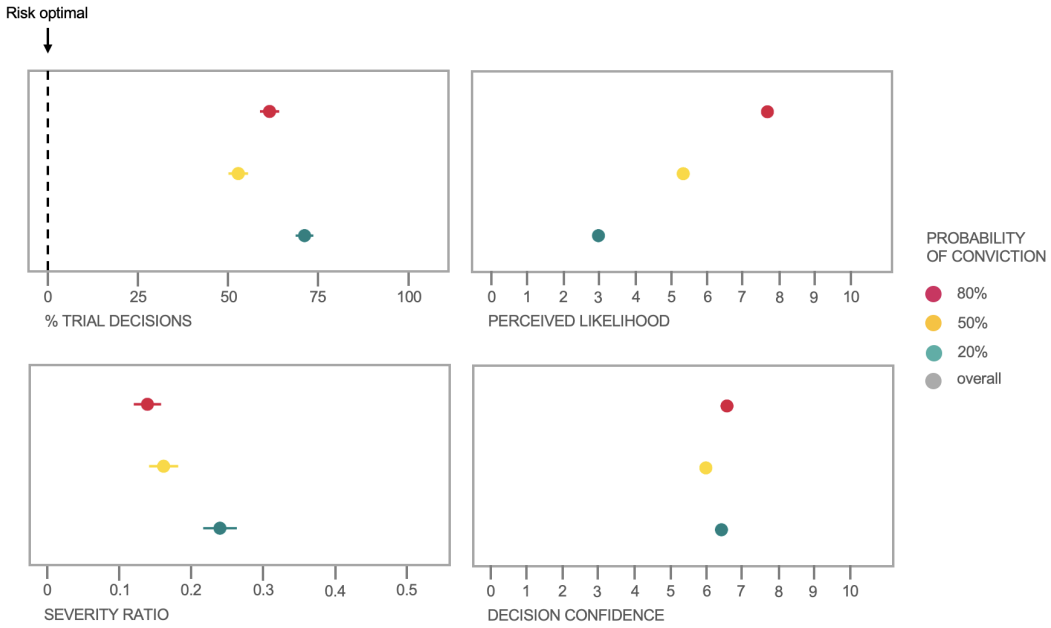

Figure 3: Experiment 1 mean and 95% CI percentage of the trial (risk optimal) decision, perception of the likelihood of conviction at trial, the ratio of the severity of trial to plea sentence, decision confidence. Model $Exp1_{confidence}$, Model $Exp1_{likelihood}$ and Model $Exp1_{severity}$ revealed a significant effect of probability of conviction.

4.6 Discussion

Our findings indicate that participants are more likely to choose the plea bargain at a 20% probability of conviction at trial compared to 50% and 80%. However, participants were more likely to choose the plea at a 50% probability of conviction compared to 80%. Therefore, our findings only partly align with the existing literature on decision-making under uncertainty, which states that higher probabilities of negative outcomes typically lead to more conservative choices [10, 70]. Investigations into participants' self-reported perceived likelihood of conviction at trial and severity ratio across probability of conviction values revealed patterns consistent with risk aversion.

The unexpected result that a 50% probability of conviction led to more conservative choices—where participants were more likely to accept the plea deal—compared to an 80% probability of conviction warrants further exploration. Analyzing participants' self-reported justifications revealed that, at the 50% probability level, they frequently relied on heuristic-based reasoning, using terms such as "chance", "coin flip", "gamble", "fifty-fifty", or "roll the dice" (see supplementary material for data). This pattern suggests that participants perceived the decision as inherently ambiguous, framing it as a high-uncertainty scenario rather than a straightforward probabilistic judgment. Ambiguity aversion—a well-documented phenomenon in decision-making research of a tendency to favor known outcomes of unknown ones—may explain this behavior [3, 45]. When faced with uncertainty about conviction outcomes, participants appeared to adopt a risk-averse approach, opting for the certainty of a plea deal over the unpredictability of a trial. This interpretation is further supported by the finding that participants were significantly less confident in their decisions when the probability of conviction was 50% compared to other probability levels. Lower confidence at this threshold suggests greater difficulty in forming a decisive expectation about trial outcomes, reinforcing the tendency toward conservative decision-making.

Participants who saw the left-skewed distribution made significantly fewer trial (risk optimal) decisions than normal distributions. These findings demonstrate that participants adjusted their decisions based on skewness. They also align with models of decision-making, which posit that people are more likely to choose the risk-averse option when presented with left-skewed distributions, where the probability of significant losses is higher [10, 70]. Although the right-skewed chart, representative of real-life trial sentence distributions, did not elicit significantly different decisions compared to the normal chart, our findings hold significant implications beyond the plea bargain scenario. Ultimately, understanding how skewness impacts decisions can guide the development of interventions that help users make more optimal choices. These insights are particularly relevant for designing decision-support systems for left-skewed distributions, such as prison sentence completion, where most inmates serve close to the full sentence, with only a few released early due to parole or other factors.

5 EXPERIMENT 2: THE EFFECT OF UNCERTAINTY & VISUALIZATION

In Experiment 1, we observed that the skewness of the data can influence decisions related to plea bargains. However, this effect is

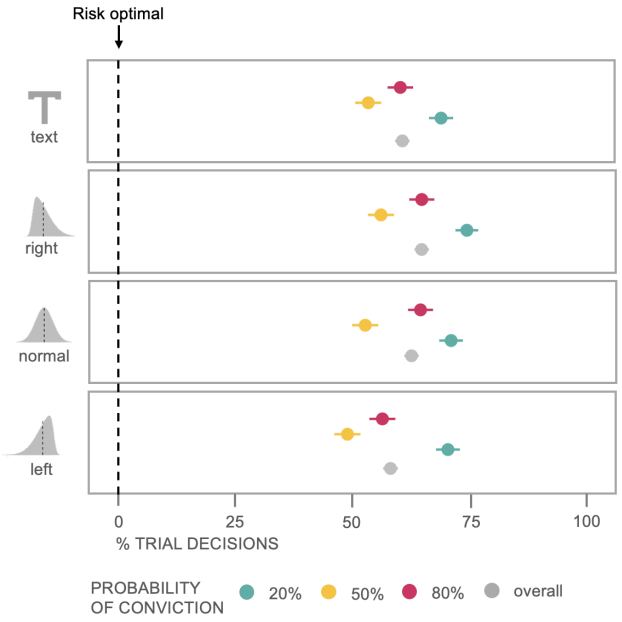

Figure 4: Experiment 1 mean and 95% CI percentage of trial, i.e., risk optimal decisions across charts and probability of conviction. Model $Exp1_{decision}$ revealed significant main effects of probability of conviction and charts where the *left* chart elicited significantly fewer trial (risk optimal) decisions compared to the *normal* chart

mostly relevant in cases of left-skewed data, i.e., where the distribution suggests a higher chance of receiving a long sentence. Despite this observation, it is important to note that such data skewness is not commonly found in sentencing data based on the available evidence; right-skewed data is more prevalent. As such, our subsequent investigations focus on Gaussian distributions due to the lack of a significant difference in outcomes when comparing normal and right-skewed distributions in Experiment 1. Additionally, using a Gaussian distribution for these subsequent investigations will allow for easier comparison with previous research. However, we do recognize that this approach may have trade-offs in identifying potential interaction effects.

In Experiment 2, the goal is to isolate the effect of data representation on decision-making. These representations include five uncertainty visualizations, a text condition showing uncertainty, and a text condition depicting only the mean. We evaluate a text condition with a deterministic mean and a text condition with uncertainty which specifies the range, mean and standard deviation, to investigate ***whether*** the presence of uncertainty impacted plea bargain decisions. This experiment also aims to examine ***how*** visualizations affect plea bargain decisions. Using methods from prior work, we evaluate participants' risk behavior to obtain a more granular measure of decision-making [6, 18]. Furthermore, **we examine the mediating effect of individual differences in numeracy, general risk propensity (GRP), and demographic factors.**

5.1 Data Presentation Conditions

Participants will be assigned to one of 6 presentation conditions between subjects: Hypothetical Outcome Plots (*HOPs*), Quantile Dotplots (*dotplot*), Probability Density Functions (*density*), Interval Plots (*interval*), text communicating uncertainty ($text_{uncertainty}$) and text communicating only the mean ($text_{mean}$) (see figure 1). In all the visualizations, with the exception of *HOPs*, the mean was explicitly represented as the point of maximum height. In the instructions (see supplementary material), participants were given a short explanation of how to interpret the visualizations. In the main task, for each decision sheet, participants were reminded of the drug crime scenario and saw the prompt *"Your attorney shows you this graph showing the jail time served by convicted defendants with similar charges as you:"* followed by their assigned presentation condition and decision sheet (see figure 5). No additional information about the visualizations was presented.

- $text_{mean}$: In order to examine whether participants make better decisions using uncertainty information, we used a mean-only condition using text as our control. This condition was intentionally deterministic to examine whether the presence of uncertainty impacted decisions.
- $text_{uncertainty}$: To investigate the potential benefit of uncertainty visualizations, we used text conditions that communicated the mean, range, and standard deviation.
- *interval*: These communicate a 95% prediction interval through the longer, thinner lines and a 66% prediction interval using the thicker, shorter lines.[4]
- *density*: As mentioned in Experiment 1, PDFs are one of the most extensively used visualizations to communicate uncertainty.
- *dotplot*: Prior work has found that quantile dotplots lead to the least bias in magnitude estimation [38]. We showed users dotplots where each of the 50 dots represented a 2% chance of a corresponding possible outcome on the x-axis.
- *HOPs*: We used animated sequences of strips representing 50 outcomes sampled from a distribution of possible outcomes (matching the data shown in quantile dotplots). Animations were rendered at 2.5 frames per second with no animated transitions (i.e., tweening or fading) between frames, looping every 20 seconds.

5.2 Participants

Participants were recruited from the crowd-sourcing platform Prolific. In this experiment, we added screeners such that the demographic of our participant pool mirrored demographics from the Cook County defendant population. We recruited a custom quota sample with 76% Black/African American, 13% White/ Caucasian, 10% Hispanic/ Latino, 1% Other [34]. The experiment was between subjects where each participant saw the trial sentence using one of the 6 uncertainty visualization conditions. We estimated the task to last about 15 minutes. Participants received a base rate of $12/h. In each of their lottery sheets, a row was picked at random and their bonus depended on their choice for that row.

[4]These charts were created using *ggdist*: https://mjskay.github.io/ggdist/

5.3 Tasks & Procedures

Before the main task, participants were shown a brief explanation of the uncertainty visualization if they were assigned one of the visualization conditions, followed by instructions and a practice round with their assigned presentation condition. The main task consisted of three decision sheets plus an attention check sheet, each containing 20 options from which participants chose when to accept a plea deal (values ranged from 0.5 to 10 years) and when to opt for the trial (see figure 5). The expected value of the trial remained constant across all decision sheets, but the probability of being convicted at trial varied, with $p = \{0.2, 0.5, 0.8\}$, corresponding to average potential trial sentences $E(S) = \{25, 10, 6.25\}$ years if convicted. In the attention check sheet, the probability of conviction at trial was 100% for an average sentence of 10 years - so participants should choose the plea if they are paying attention to the task. After each decision, participants estimated on a Likert scale their perceived likelihood of conviction *("What was your likelihood of conviction at trial?")* and decision confidence *("How confident are you in your decision")*.

We randomly selected a line from the 20 choices in the decision sheet. If participants chose the trial, we simulated the trial outcome and sentence length using the trial data. If they chose the plea, they obtained the guaranteed plea sentence. Participants then saw their simulated trial outcome and sentence length. They were given an initial bonus of $7 for this task. For each year of their sentence, $0.01 was deducted from their bonus. At the end of the trials, participants completed a numeracy survey from [25], consisting of a total of 12 items, with 11 items from one scale [44] and one additional item from another [63]. They also completed a General Risk Propensity (GRP) Assessment where they answered 8 questions to assess their general risk propensity [72] and a demographic survey. At the end of the study, the participant saw a bonus sheet that included their total jail time throughout the study and the remaining bonus.

5.4 Evaluating Risk Behavior

Experiment 1 measured participants' responses to a binary choice about taking the plea or going to trial. While this gave us insight into their risk behavior, other frameworks exist to measure decision-making with more granularity. Relative Risk Premia (RRP) is a concept from economics and finance that quantifies an individual's risk aversion when making a series of decisions involving uncertain outcomes [18]. RRP allows us to evaluate not only whether a decision-maker is risk-seeking, risk-neutral, or risk-averse but also the degree to which they exhibit these tendencies. In this work, we refer to a risk-neutral decision as being **risk optimal**, per behavioral economics standards.

$$RRP = \frac{(pE(S) - ce)}{|pE(S)|}$$

where $pE(S)$ denotes the expected value of the trial outcome, which is set to 5 years, and ce denotes the certainty equivalent of the trial. We calculate the trial's certainty as equivalent to the average of the last smallest plea deal the participant selected on the sheet and the subsequent plea deal offered. $RRP > 0$ **indicates risk seeking behavior,** $RRP < 0$ **implies risk averse behavior and** $RRP = 0$ **suggests risk optimality.**

5.5 Analysis

We conducted a series of linear mixed-effects models (LMEs) to evaluate the influence of uncertainty, visualization, and individual differences on *RRP*. Fixed effects included the probability of conviction (*p*), visualization (*format*), individual differences in numeracy (*numeracyScore*) and general risk propensity (*riskScore*), demographic factors (*Gender*, *Age*, *Education* and *Ethnicity*) as well as whether the participants have been accused or convicted of a crime (*priorCrime*) or offered a plea bargain (*priorBargain*). Participant *id* was included as a random intercept to account for repeated measures within subjects. We conducted similar full mixed effects models to investigate the impact of all our predictors on the perception of the likelihood of conviction (*perceptionOfLikelihood*) and decision confidence (*decisionConfidence*). The outcome variables were transformed using a *boxcox* power transformation to improve normality. In this section, we reported the findings from our three main models as well as any follow-up models to examine interaction effects. Across all the LMEs, the reference level for the representation format was set to $text_{uncertainty}$ to enable comparisons of the presence of uncertainty and uncertainty visualizations.

Our main models to examine the effect of visualization and individual differences on decision-making were:

- $Exp2_{decision}$: $RRP \sim \text{lmer}(p + format + numeracyScore + riskScore + Gender + Age + Education + Ethnicity + priorCrime + priorBargain + (1 \mid id))$
- $Exp2_{likelihood}$: $perceptionOfLikelihood \sim \text{lmer}(p + format + numeracyScore + riskScore + Gender + Age + Education + Ethnicity + priorCrime + priorBargain + (1 \mid id))$
- $Exp2_{confidence}$: $decisionConfidence \sim \text{lmer}(p + format + numeracyScore + riskScore + Gender + Age + Education + Ethnicity + priorCrime + priorBargain + (1 \mid id))$

where:

$decision$	$\in [-1.05 \ldots 0.95]$; RRP ranging from always choosing the plea to always choosing the trial
$decisionConfidence$	$\in [0 \ldots 10]$; confidence in the decision
$perceptionOfLikelihood$	$\in [0 \ldots 10]$; the perceived likelihood of going to trial

5.6 Results

Probability of Conviction. The median *RRP* across 20%, 50%, and 80% probability of conviction was 0.75 (mean = 0.46, sd = 0.61), 0.55 (mean = 0.41, sd = 0.54), and 0.35 (mean = 0.21, sd = 0.62) respectively. Model $Exp2_{decision}$ revealed that probability of conviction had a significant effect on risk behavior, where a 50% (Estimate = -0.08, SE = 0.04, z = -2.12, $p < 0.05$) and 80% (Estimate = -0.29, SE = 0.04, z = -7.37, $p < 0.001$) probability of conviction elicited a lower *RRP* i.e., less risk-seeking (more risk optimal) behavior compared to the baseline 20%. The probability of conviction revealed no significant impact on the perception of the likelihood of conviction or decision confidence. *Consistent with Experiment 1, our findings revealed that participants made less risk-seeking decisions when the probability of conviction at trial was higher.*

Uncertainty vs. Point Estimate. The median *RRP* when participants saw $text_{mean}$ and $text_{uncertainty}$ were 0.45 (mean = 0.30,

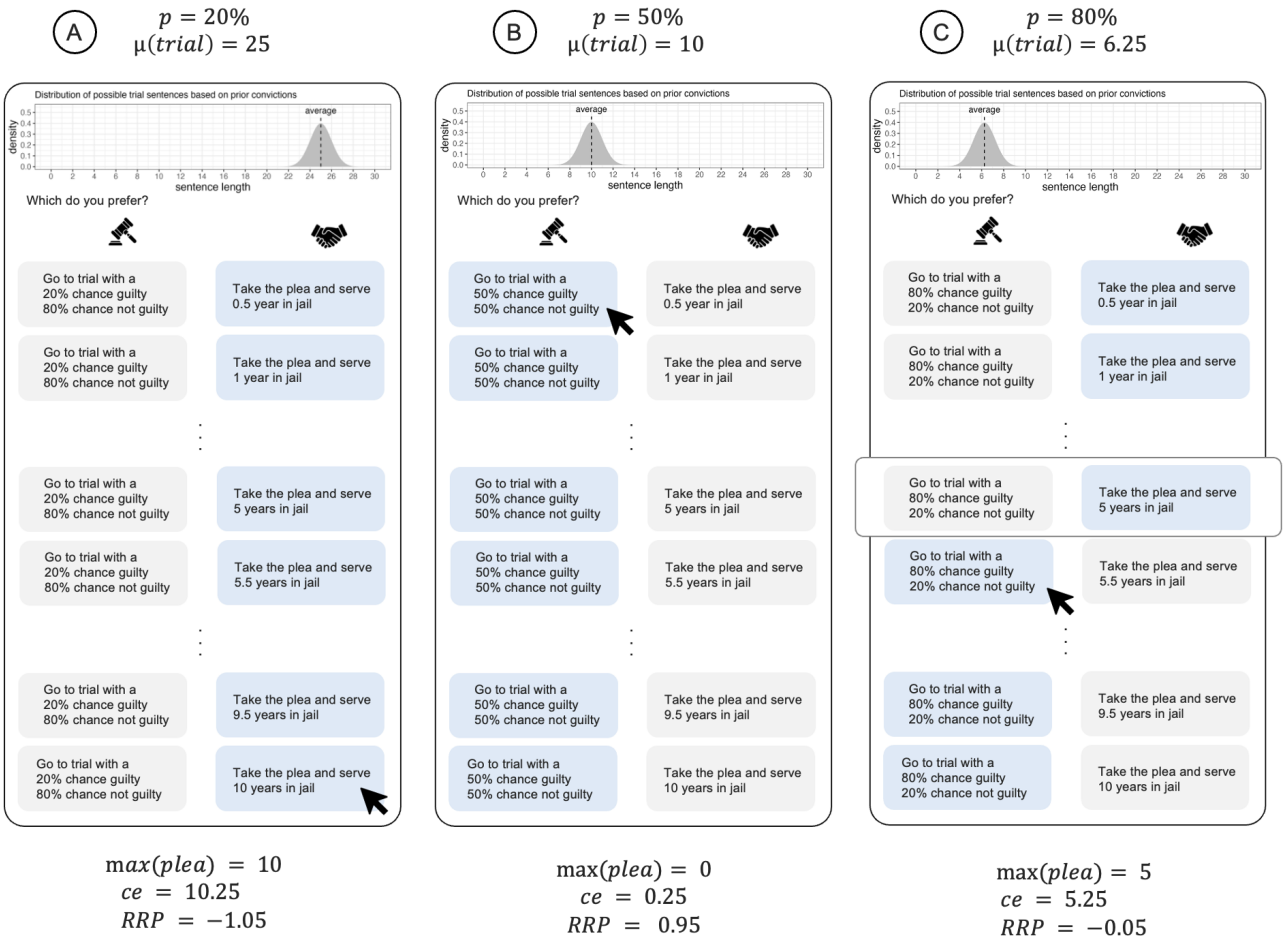

Figure 5: Examples of decision sheets with *density*. In each decision sheet, participants were presented with 20 choices about whether to go to trial or accept a plea deal (incremented by 0.5 years for each choice). Participants only needed to select their last plea decision or first trial decision, and the rest of the choices were selected accordingly. A) The most risk-averse decision for 20% probability of conviction B) The most risk-seeking decision for 50% probability of conviction C) Risk optimal decision for 80% probability. The circled selection where *E(trial)* = *E(plea)* yields an *RRP* of either -0.05 (plea decision) or -0.05 (trial decision), the two most risk optimal decisions.

sd = 0.62) and 0.45 (mean = 0.29, sd = 0.60) respectively. Model $Exp2_{decision}$ revealed no significant change in *RRP*, perception of the likelihood of conviction, and decision confidence when participants used $text_{mean}$ compared to the baseline $text_{uncertainty}$. *Our findings indicate that conveying uncertainty within text did not significantly influence decision-making compared to presenting only the mean.*

Uncertainty Visualizations. The median *RRP* when participants saw *density*, *dotplot*, *HOPs*, *interval* and $text_{uncertainty}$ was 0.55 (mean = 0.39, sd = 0.58), 0.55 (mean = 0.40, sd = 0.55), 0.55 (mean = 0.36, sd = 0.64), 0.65 (mean = 0.44, sd = 0.55) and 0.45 (mean = 0.29, sd = 0.60) respectively. Model $Exp2_{decision}$ revealed no significant change in *RRP*, perception of the likelihood of conviction at trial, or decision confidence when using charts compared to the baseline $text_{uncertainty}$. *Our findings revealed no significant impact of uncertainty visualizations on decision-making compared to text. However, text elicited marginally more risk-neutral decisions than charts when communicating the mean alone and uncertainty.*

Individual Differences in Numeracy & General Risk Propensity. Across participants, the median numeracy score was 9 out of 12, and the median risk score was 28 out of 48. Our main *RRP* model ($Exp2_{decision}$) showed that a high numeracy score could cause significantly less risk-seeking (i.e., more risk optimal) decisions (Estimate = -0.03, SE = 0.01, z = -2.71, $p < 0.05$). Our perception model ($Exp2_{likelihood}$) also showed a significant impact of numeracy score on the perception of the likelihood of conviction (Estimate = -0.08, SE = 0.04, z = -2.20, $p < 0.05$), where participants with high numeracy scores tended to perceive their likelihood of conviction

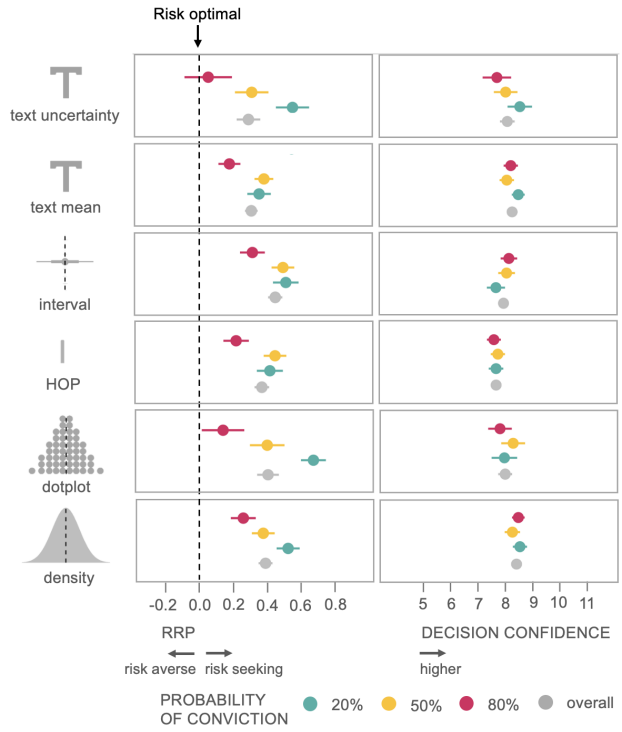

Figure 6: Experiment 2 mean and 95% CI RRP and decision confidence across probability of conviction and format. Our findings revealed no significant effect of format on decision-making.

as lower than participants with low numeracy scores. The risk score effect on *RRP* was insignificant. However, model $Exp2_{confidence}$ revealed that a high-risk score could cause significantly higher confidence in decisions (Estimate = 0.17, SE = 0.06, z = 2.76, $p <$ 0.01). Risk score also significantly impacted the perception of the likelihood of conviction (Estimate = 0.03, SE = 0.01, z = 3.02, $p <$ 0.01). We conducted follow-up mixed effects models to examine interaction effects between probability of conviction, visualization, numeracy score, and risk score on decisions but found no significant effects. A Pearson test revealed a small but significant inverse relationship between numeracy score and risk score, showing that participants with higher numeracy tend to be more risk averse than those with lower numeracy (r = -0.097, $t(1045)$ = -3.17, $p < 0.05$). *Our main takeaway from these findings is that participants with high numeracy make more optimal plea bargain decisions, and participants with a high general risk propensity tend to be more confident in their decisions.*

Demographics & Precedents. Our sample consisted of 308 males, 38 females, 2 non-binary, and 1 "prefer not to say". Five reported to have less than a high school education, 112 reported a high school education, 148 a bachelor's degree, 73 a master's degree, and 11 a doctoral degree. 260 were Black/African American, 35 were Latino/Hispanic, 46 were White/Caucasian, and 3 reported Other. The median age of our participants was 39 years old. 24 participants reported having been accused or convicted of a drug crime in the past, while 320 reported no and 5 preferred not to say. 35 reported having experience with being offered a plea bargain, while 310 reported no, and 4 preferred not to say. We conducted post-hoc analyses to examine the association of demographic factors and precedents on decisions and decision strategies. *Our findings reveal a lack of evidence for an association between education, race, age, or precedent on decision-making.*

5.7 Discussion

Consistent with findings from Experiment 1, we found that the probability of conviction at trial had a significant effect on plea bargain decisions, where 50% and 80% probability of conviction elicited more risk-neutral decisions compared to 20% probability of conviction. A key goal of our study was to investigate whether presenting uncertainty could improve decision-making. Specifically, we sought to determine if providing participants with both the mean and the associated uncertainty, as opposed to just the mean, would result in better decisions. Our findings, however, revealed no significant differences in decision-making when participants were presented with text conveying the mean alone or uncertainty. This suggests that simply adding uncertainty information did not have a noticeable impact on how participants approached their decisions. This lack of significant difference may indicate that, in this context, participants relied primarily on the central tendency (the mean) when making decisions, potentially disregarding or underutilizing the additional uncertainty information. Kale et al. found that adding means to uncertainty visualizations has minor biasing effects on magnitude estimation and decision-making, consistent with discounting uncertainty [39]. Although it falls outside the scope of this study, our findings raise important questions about how uncertainty is perceived and processed in decision-making, which future research should explore.

This study also aimed to investigate the impact of uncertainty visualizations on plea bargain decision-making. The finding that text representations (conveying both mean and uncertainty) elicited marginally more risk-neutral decisions than charts on average may interest both designers and researchers. Given the significant impact of plea bargaining scenarios, it may be beneficial to consider the small advantages a design choice can offer. However, most empirical studies in visualization research primarily focus on comparing the effect of different charts on task performance, overlooking the potential of textual or numerical representations as a viable or even superior alternative for data communication. While it is commonly believed that visualizations make it easier to understand data compared to text, our findings contribute to a growing body of work indicating that text can be as effective as, and sometimes better than, visual representation. Prior work has shown evidence that text can be better than charts in some tasks, such as Bayesian reasoning [47, 50, 51]. Several researchers have also examined how text can be used to augment the efficacy of visualizations. For example, Stokes et al. have examined the effect of combining text and visualization and advocated for more text evaluation in visualization research [69].

Since the demographics of defendants do not typically reflect the demographics of the general population, our work also focuses on

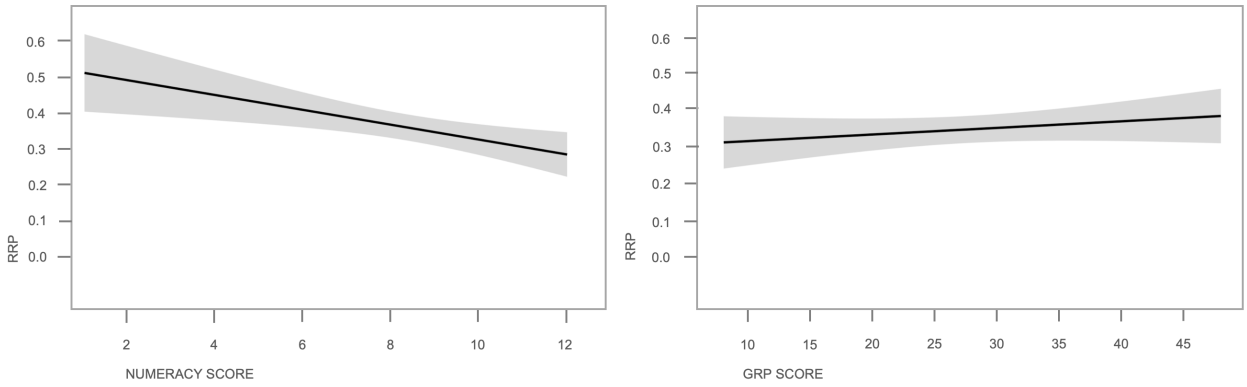

Figure 7: Experiment 2 linear relationship between numeracy scores and RRP (top) and GRP scores and RRP (bottom). Our findings revealed that participants with high numeracy made more optimal decisions, and those with high GRP were more confident in their decisions.

examining the impact of numeracy, general risk propensity, demographic factors, and precedents on plea bargain decision-making. Participants with higher numeracy perceived their likelihood of conviction at trial as lower than participants with lower numeracy and made significantly more optimal decisions. This finding underscores the critical role of numeracy in effective decision-making. In situations like plea bargains, where decisions often hinge on understanding probabilities, individuals with higher numeracy are better equipped to make informed, rational choices. This suggests that presentation formats aimed at supporting comprehension for low-numeracy individuals could also enhance the quality of their decisions. Conversely, initiatives aimed at enhancing numeracy in the general population could significantly improve decision quality in contexts requiring critical judgment. Our findings also revealed that individual differences in general risk propensity affected decision confidence, where participants with higher GRP reported significantly higher confidence compared to those with lower GRP. While decision confidence in the context of plea bargain decisions has been understudied, high decision confidence in other contexts has been shown to be correlated with increased commitment to the decision, greater satisfaction with the outcome, and reduced likelihood of decision reversal [64].

6 EXPERIMENT 3: THE EFFECT OF TIME PRESSURE

During negotiations, the prosecutor might offer a plea deal that expires within a short time frame [13]. This tactic can pressure defendants into making decisions potentially without adequate time to consult with their attorney or family. Many applications beyond the plea bargain scenario require time-sensitive decision-making, including emergency medicine, military, and defense, space mission control, or disaster response. However, there is limited research on the impact of time pressure on visualization-assisted decision-making. In this experiment, we conduct a similar task to Experiment 2 but add a time constraint to the plea bargain decisions. We aim to examine how time pressure affects plea bargain decision-making with various uncertainty visualizations and how individual differences play a role.

6.1 Tasks & Procedures

In this experiment, we used the same task as Experiment 2 but this time, participants had 17 seconds to complete each decision sheet. Given the high variance in participants' average time taken per trial, we chose to deviate from the time pressure metrics used in previous research [42]. In this experiment, the time limit was set to 60% of the mean time taken per trial in Experiment 2 (mean = 28, sd = 19.9), i.e., 17 seconds. If they fail to make a decision before the timer runs out, they automatically go to trial, where their trial outcome and sentence are simulated. At the end of the experiment, participants completed a numeracy test, a GRP test, and a NASA-TLX questionnaire.

6.2 Participants

We recruited 347 participants on Prolific and set the same screening criteria as Experiment 2, such that our participant pool mirrors the defendant population. After filtering out participants who did not finish the survey or took less than 3 seconds on average to answer decision questions, we were left with 336 participants.

6.3 Analysis

In Experiment 3, after data exclusions due to attention checks and data quality, 336 participants remained. However, 1 participant failed to make a decision before the timer was up across all three trials, and 59 participants failed to make a decision before the timer was up for at least 1 out of 3 trials. Due to missing data, we removed these trials from our dataset prior to all data analysis involving experiment 3. Therefore, our final dataset consists of 335 participants across a variable number of trials (1, 2 or 3). We conducted a linear mixed-effect model to examine the effect of time pressure (*timePressure*) and visualization (*format*) on *RRP*, with random intercepts for each participant *id* to account for repeated measures. The outcome variables were transformed using a *boxcox* power

transformation to improve normality. Across all the LMEs, the reference level for the representation format was set to $text_{uncertainty}$, allowing comparisons of the presence of uncertainty as well as uncertainty visualizations. Our models to examine the effect of time pressure were:

- $Exp3_{time}$: $RRP \sim \text{lmer}(timePressure + (1 \mid \text{id}))$
- $Exp3_{vistime}$: $RRP \sim \text{lmer}(p * format * timePressure + (1 \mid \text{id}))$

where:

where:

$timePressure \in \{0, 1\}$; unlimited time or limited time

In this section, we report the findings from three main models as well as any follow-up models to examine interaction effects. Our main models were consistent with the models conducted in Experiment 2 (see section 5.5):

- $Exp3_{decision}$: $RRP \sim \text{lmer}(p + format + numeracyScore + riskScore + Gender + Age + Education + Ethnicity + priorCrime + priorBargain + (1 \mid \text{id}))$
- $Exp3_{likelihood}$: $PerceptionOfLikelihood \sim \text{lmer}(p + format + numeracyScore + riskScore + Gender + Age + Education + Ethnicity + priorCrime + priorBargain + (1 \mid \text{id}))$
- $Exp3_{confidence}$: $DecisionConfidence \sim \text{lmer}(p + format + numeracyScore + riskScore + Gender + Age + Education + Ethnicity + priorCrime + priorBargain + (1 \mid \text{id}))$

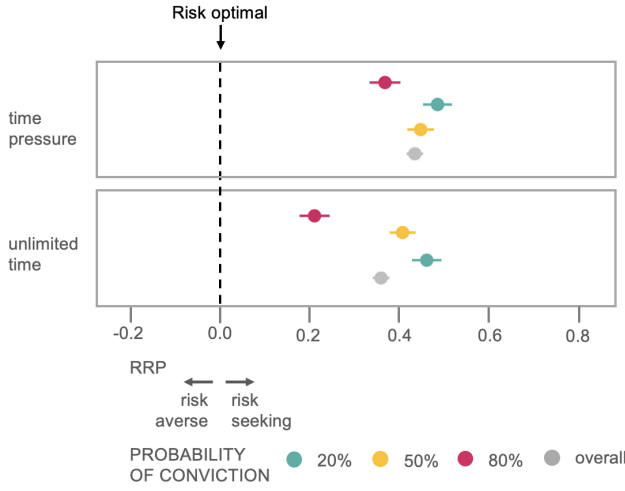

Figure 8: Time pressure (Experiment 3) versus no time pressure (Experiment 2) RRP. Participants made significantly less risk optimal decisions when subject to time pressure when the probability of conviction at trial was 80%.

6.4 Results

The Effect of Time Pressure on Decisions. We combined datasets from Experiment 2 (unlimited time) and Experiment 3 (time pressure) to examine the effect of time pressure on decisions and their interaction with probability of conviction and visualization. We conducted a linear mixed-effect model $Exp3_{time}$ to examine the

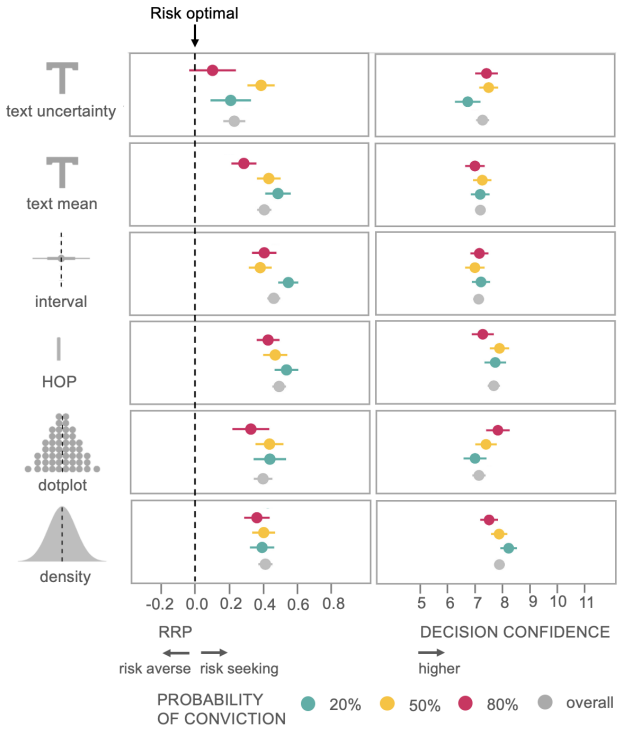

Figure 9: Experiment 3 mean and 95% CI RRP and decision confidence across probability of conviction and visualization. Our findings revealed that $text_{uncertainty}$ elicited more optimal decisions compared to all other formats.

effect of time pressure on *RRP* and found that the presence of time pressure caused participants to make significantly more risk-seeking behavior (Estimate = 0.09, SE = 0.04, z = 2.41, $p < 0.05$). Model $Exp3_{vistime}$ revealed that time pressure had a significant interaction effect with 50% (Estimate = 0.50, SE = 0.22, z = 2.34, $p < 0.05$) and 80% probability of conviction, again causing more risk-seeking behavior (Estimate = 0.45, SE = 0.22, z = 2.04, $p < 0.05$). While we found no main effect of representation format, we found interaction effects of time pressure with *HOP* (Estimate = 0.55, SE = 0.22, z =2.50, $p < 0.05$), *interval* (Estimate = 0.46, SE = 0.22, z = 2.05, $p < 0.05$) and $text_{mean}$ (Estimate = 0.56, SE = 0.22, z = 2.56, $p < 0.05$), where the added time pressure elicited significantly more risk-seeking (less risk optimal) behavior across these charts. *Our findings revealed that time pressure resulted in participants making less risk-neutral decisions, especially when using HOPs, interval, and* $text_{mean}$.

Probability of Conviction. The median *RRP* across 20%, 50%, and 80% probability of conviction was 0.75, 0.65, and 0.55, respectively. Model $Exp3_{decision}$ revealed that under time pressure, an 80% (Estimate = -0.16, SE = 0.05, z = -3.45, $p < 0.001$) probability of conviction at trial elicited more risk-averse behavior compared to the baseline of 20% probability of conviction. We found no significant main effect of the probability of conviction on the perception of the likelihood of conviction or decision confidence. *Our findings revealed that similarly to Experiments 1 and 2, a higher probability*

of conviction at trial caused more risk-averse behavior under time pressure conditions.

Uncertainty vs. Point Estimate The median *RRP* when participants saw $text_{uncertainty}$ and $text_{mean}$ was 0.45 (mean = 0.25, sd = 0.63) and 0.65 (mean = 0.43, sd = 0.57) respectively. Model $Exp3_{decision}$ revealed a significant increase in RRP when participants used $text_{mean}$ in relation to the reference level $text_{uncertainty}$ (Estimate = 0.23, SE = 0.11, z = 2.09, $p < 0.05$). We found no significant effect of uncertainty on the perception of the likelihood of conviction at trial and decision confidence. *We can conclude that under time pressure, presenting uncertainty elicited decisions closer to risk optimality compared to presenting only the mean.*

Uncertainty Visualizations. The median *RRP* when participants saw $text_{uncertainty}$, *density*, *interval*, *dotplot*, and *HOP* was 0.45 (mean = 0.25, sd = 0.63), 0.65 (mean = 0.42, sd = 0.56), 0.65 (mean = 0.48, sd = 0.53), 0.75 (mean = 0.42, sd = 0.60), and 0.75 (mean = 0.52, sd = 0.52) respectively. We found a significant effect of *density* (Estimate = 0.22, SE = 0.11, z = 2.04, $p < 0.05$), *dotplot* (Estimate = 0.23, SE = 0.12, z = 1.97, $p < 0.05$), *HOP* (Estimate = 0.37, SE = 0.11, z = 3.35, $p < 0.001$) and *interval* (Estimate = 0.32, SE = 0.11, z = 2.97, $p < 0.01$) on RRP, where they elicited significantly more risk-seeking behavior compared to the baseline $text_{uncertainty}$. We found no significant effect of visualization on the perception of likelihood or decision confidence. *Our findings reveal that $text_{uncertainty}$ elicited more risk-neutral plea bargain decisions compared to charts.*

Individual Differences in Numeracy & General Risk Propensity. The median numeracy score across our participants was 9 out of 12, and the median risk score was 28 out of 48. Model $Exp3_{decision}$ revealed no significant effect of numeracy score on RRP, perception of likelihood, or decision confidence. We conducted a follow-up mixed effects linear model to examine the interaction effects between probability of conviction, visualization, numeracy score, and risk score. Our model revealed that when using *HOPs*, higher numeracy scores elicited significantly less risk-seeking (more risk optimal) behavior (Estimate = -0.61, SE = 0.26, z = -2.37, $p < 0.05$). *This finding might suggest that HOPs were harder to understand compared to other representations.* Model $Exp3_{confidence}$ revealed that risk score significantly impacted decision confidence, where higher risk scores were associated with higher decision confidence (Estimate = 0.10, SE = 0.02, z = 4.63, $p < 0.001$). *This finding that participants with a high general risk propensity tend to be more confident in their decisions is consistent with Experiment 2.*

Demographic Factors & Precedents. Our sample consisted of 296 male, 37 female, and 2 non-binary participants. Two participants reported having less than a high school education, 109 reported a high school education, 165 a Bachelor's degree, 51 a master's degree, and 8 a doctoral degree. Fourteen participants reported having been accused or convicted of a crime, and 16 reported having previously been offered a plea bargain. The median age of our participants was 35 years old. Model $Exp3_{decision}$ revealed that older participants made significantly more risk-seeking decisions compared to younger participants (Estimate = 0.009, SE = 0.01, z = 3.50, $p < 0.001$). Model $Exp3_{confidence}$ also revealed that older participants reported significantly higher decision confidence compared to younger participants (Estimate = 0.05, SE = 0.02, z = 2.27, $p < 0.05$). A follow-up Pearson correlation revealed a significant small negative correlation between age and numeracy score, suggesting that older participants had lower levels of numeracy compared to younger individuals ($r\,(938) = -0.11$, $p < 0.05$). *Older participants tended to have lower numeracy levels than younger participants, which could explain why they made less risk-neutral decisions.*

Model $Exp3_{decision}$ also revealed that White/Caucasian participants made significantly more risk-averse decisions compared to Black/African American participants (Estimate = -0.22, SE = 0.09, z = -2.58, $p < 0.05$). We conducted follow-up analyses to examine whether Ethnicity is a predictor of numeracy or general risk propensity. White/Caucasian participants had a median numeracy score of 26 (mean = 24.98, sd = 9.81), and Black/African American participants had a median numeracy score of 30 (mean = 28.49, sd = 10.35). We conducted a Kruskal-Wallis test and found that Ethnicity revealed a significant association with general risk propensity scores (χ^2 = 59.517, $p < 0.001$). Follow-up pairwise Wilcoxon tests with a Bonferroni adjusted alpha revealed significant differences in risk score among all pairs of Ethnicity groups except "Other". White/Caucasian participants had a median numeracy score of 10 (mean = 10.17, sd = 1.06), and Black/African American participants had a median numeracy score of 9 (mean = 8.27, sd = 2.54). We conducted a Kruskal-Wallis test and found that Ethnicity demonstrated a significant association with numeracy score (χ^2 = 90.822, $p < 0.001$). We conducted follow-up pairwise Wilcoxon tests with a Bonferroni adjusted alpha and found significant differences in numeracy scores between all pairs of Ethnicities except White and Other, where Black/African American participants had lower numeracy than other ethnicities. Model $Exp3_{confidence}$ also revealed that Latino/Hispanic (Estimate = -1.75, SE = 0.76, z = -2.30, $p < 0.05$), White/Caucasian (Estimate = -1.44, SE = 0.68, z = -2.12, $p < 0.05$) and Other (Estimate = -6.79, SE = 2.38, z = -2.85, $p < 0.001$) participants reported significantly lower confidence in their decisions compared to Black/African American participants. The association between gender, education, precedent, and decisions was not significant. *Our findings revealed that numeracy and general risk propensity levels vary amongst ethnic groups, where Black/ African American participants exhibited lower numeracy and higher general risk propensity than White participants. This could explain why Black/ African American participants made significantly fewer risk-neutral decisions than White participants.*

6.5 Discussion

We found that time pressure led participants to make significantly less risk-neutral (i.e., more risk-seeking) decisions than decisions made under no time constraints. This effect was especially salient when the likelihood of being guilty at trial was higher, at 80% probability of conviction.

Under time pressure conditions, we found that a 50% and 80% probability of conviction elicited more risk-optimal decisions compared to 20%. This finding that a higher probability of conviction at trial elicits more risk-averse decisions is consistent with Experiments 1 and 2. Contrary to Experiment 2, we found that communicating uncertainty via text elicited more optimal decisions than communicating only the mean. Another goal of this experiment was to investigate whether uncertainty visualizations impacted decision-making compared to text under time-pressure conditions.

Our findings revealed that text was the superior format since it elicited more optimal decisions than all charts.

When investigating the mediating effect of individual differences on numeracy, general risk propensity, demographic factors, and precedents on decision-making under time pressure, we found that participants with higher general risk propensity reported significantly higher confidence in their decisions – a finding consistent with Experiment 2. Our analyses revealed an interaction where when using *HOP*s, a higher risk score was associated with more risk optimal decisions, i.e., less risk seeking. This could suggest that *HOPs* are harder to use compared to other charts when time pressure is involved due to their animated nature.

While findings from Experiment 2 revealed no association between demographic factors and decision-making, we found that age and race were significantly associated with decision-making under time pressure. Older participants make significantly less optimal decisions compared to younger participants. Further analyses revealed that older participants tend to have lower numeracy levels than younger participants. We also found that Black/ African American participants made significantly less optimal decisions with higher confidence compared to other races. Further analyses suggest that this effect may be attributed to differences in numeracy and general risk propensity. Black/African American participants exhibited significantly lower numeracy levels compared to participants from other racial groups. This finding highlights broader systemic issues, such as educational inequities and socio-economic factors, that disproportionately affect certain demographic groups, leading to disparities in numeracy and, consequently, in decision-making abilities. These systemic challenges can have far-reaching implications, potentially influencing outcomes in legal, financial, and health-related decisions, where numerical literacy is critical.

7 GENERAL DISCUSSION & LESSONS LEARNED

This work was motivated by the pressing need to develop support tools to assist defendants when considering plea bargain offers, a critical juncture in the criminal justice process. While our research does not capture the entire spectrum of situational factors that might influence a defendant's decision-making process, such as emotional state, legal advice, or the strength of the prosecution's case, our studies focus on a subset of critical factors that are often overlooked in the visualization community and could have ramifications beyond plea deals. Specifically, we sought to isolate and analyze the impact of *whether* and *how* uncertainty is presented to defendants, as this can significantly shape their risk perception and, ultimately, their decisions.

Defendants often face significant pressure when deciding whether to accept a plea bargain or risk going to trial. This decision is further complicated by their personal risk tolerance and the ability to fully understand their decision's legal and statistical implications. As such, our work explores the effect of individual differences and time pressure on plea bargain decisions. Our research has implications for developing plea bargain decision aids, but our findings also revealed insights that could benefit the broader visualization community.

A Case for Text in Visualization Research. Our findings demonstrate that uncertainty communication improves decision-making compared to deterministic information and, notably, that textual uncertainty representations led to more optimal decisions than visualizations, particularly under time pressure. Traditionally, visualization research prioritizes graphical methods and we often underestimate the effectiveness of text-based approaches for conveying uncertainty. However, our results suggest that text can provide advantages in certain conditions, particularly when decisions must be made quickly.

One possible explanation for this advantage is cognitive load reduction — text may allow participants to process uncertainty directly without needing to interpret a visual encoding. Additionally, speed of interpretation may play a role, as text might facilitate a quicker understanding of probabilities. These findings suggest that visualization researchers should reconsider assumptions about the superiority of visual methods and explore hybrid or adaptive approaches that incorporate text when appropriate [8]. Therefore, we support the calls from current researchers to incorporate text as a visualization condition [29, 50, 68].

The Role of Time Pressure in Uncertainty Communication. While text-based uncertainty communication showed a marginal improvement over uncertainty visualizations under unlimited-time conditions, this effect became more pronounced under time pressure. Participants were more likely to make optimal decisions when uncertainty was presented textually rather than visually, particularly when they had limited time to process information. This suggests that the benefits of textual uncertainty representations may be particularly relevant in high-stakes or fast-paced decision-making scenarios. This finding also underscores why decision-making research — and particularly uncertainty visualization evaluation — should account for the role of time constraints, as many real-world decision-making scenarios (e.g., legal choices and emergency response) involve significant time pressure.

Toward a Better Understanding of Demographics & Visualizations. Systemic disparities in the criminal justice system disproportionately impact minority groups and less-educated defendants, often shaping how individuals navigate legal decisions. To ensure that decision-support tools are both effective and accessible across diverse populations, we aimed to account for these demographic differences.

Our investigations into the role of demographic factors in decision-making found that they were associated with decision outcomes, reinforcing the need for equitable approaches in visualization research. However, while some demographic patterns emerged, we observed that these same factors were also associated with variations in numeracy and general risk propensity. This overlap makes it challenging to determine whether differences in decision-making stem primarily from demographic background, cognitive abilities, or broader societal disparities.

These findings also highlight the need for deeper research into how demographic factors shape visualization comprehension and decision-making. Future studies should carefully control for potential confounding variables, ensuring that insights into equitable design are not conflated with individual differences in numeracy, risk perception, or experience.

Decision Framing Can Impact Observations. Our investigation of plea bargain decisions revealed some surprising effects of methodological design and question framing. In particular, we found that, in Experiment 1, where participants were asked to make a binary choice, they were less likely to choose the trial when the probability of conviction was 50% compared to 80%. Analyses of decision confidence and participant justifications in Experiment 1 suggested that the 50% conviction probability was perceived as a coin flip, leading to heuristic-based decision-making and more conservative choices. This interpretation prompted us to refine the study design in Experiments 2 and 3 to allow for a more granular analysis of participants' decision-making processes. The modifications were necessary to account for the unintended effects observed in Experiment 1's binary-choice format. Instead of offering a simple binary choice between accepting a plea deal or going to trial, Experiments 2 and 3 presented participants with 20 different plea deal options. This transformation shifted the task from a straightforward accept-or-reject decision to a more nuanced judgment: **at what point does the plea deal become unfavorable enough to justify risking the uncertain outcome of a trial?**

Our results from Experiments 2 and 3 followed the expected risk aversion pattern, where participants were less likely to choose the trial when the probability of conviction was larger. By introducing this expanded choice set, Experiment 2 minimized the influence of perceived binary uncertainty and allowed for a more precise measurement of risk-related preferences. Since participants in this design made 20 choices per decision sheet with varying plea sentences, the perceived gap between plea and trial outcomes was more systematically controlled when computing the Relative Risk Premia (RRP). As a result, we observed the expected trend; participants were more conservative (i.e., more likely to accept a plea) when the probability of conviction was 80% compared to 50%. This finding underscores how methodological frameworks shape decision-making by influencing how participants perceive and approach a given task. The observed differences highlight the importance of careful experimental design in studies of legal decision-making and risk perception. Researchers must balance methodological rigor with ecological validity to ensure that findings accurately reflect the underlying cognitive processes being investigated.

8 LIMITATIONS

Our controlled experimental design enabled us to isolate the impact of visualization on decision-making while considering some key individual differences and contextual factors relevant to the plea bargain scenario. However, our experiments do not fully capture the complexities of real-life legal decisions. Firstly, participants in our study made decisions based on hypothetical scenarios, which may not elicit the same level of emotional and cognitive engagement as actual plea bargain situations. The time pressure condition in Experiment 3 was designed to simulate high-stakes decision-making, but it is not entirely analogous to real-world scenarios where defendants facing *exploding offers* may have only a few hours, rather than seconds, to make their decisions [19, 57, 73]. Moreover, our experiments may not accurately capture the nuanced pressures defendants face in real legal settings, such as pressure from their attorneys. Finally, while the choice of mixed-design experiments ensures the adequate exploration of the effect of probabilities, visualization, and individual and contextual factors, it requires participants to make a series of decisions, whereas, in real life, they would only make one.

Moreover, we acknowledge that personal experience with criminal justice can shape decision-making. While we asked participants whether they had previously been accused or convicted of a crime or offered a plea bargain, the number of participants who answered "yes" was too small to observe the effect of personal experience. Altogether, our findings offer valuable insights into how uncertainty visualization might influence plea bargain decisions. While our study provides strong evidence within specific conditions of our study, further research is needed to explore how these effects translate to more complex, real-world legal contexts.

9 FUTURE DIRECTIONS

The finding that text led to more optimal decisions than charts calls for a deeper examination of participants' familiarity with charts and visualization literacy and how these factors influence decision-making under time pressure. While prior research has explored the design of glanceable visualizations [15], there is a lack of research on how visualizations can impact decision-making under time pressure across other scenarios. Our work also highlights how numeracy, general risk propensity (GRP), and demographic factors such as age and ethnicity can impact decisions. While we hesitate to generalize the impact of demographic factors due to their associations with cognitive abilities. Future research should conduct more investigations to better isolate these effects and more effectively support minority populations.

There are also opportunities for future work to examine other aspects of plea bargaining and how visualization can support decision-making. Future work could examine how the advice of an attorney can affect participants' decisions and whether visualizations affect those decisions. Moreover, future work could conduct qualitative analyses to better understand how people are making decisions and to what extent they are based on data or personal experiences.

In the past few years, there have been various calls to diversify evaluation metrics in visualization research [7, 23], including the creation of the BELIV (Beyond Time and Errors: Novel Evaluation Methods for Visualization) workshop [12]. We encourage researchers to consider application-relevant individual differences and circumstantial factors such as time pressure when evaluating visualizations. Understanding these effects could lead to developing more effective decision-support tools optimized to improve outcomes in high-stakes environments.

10 CONCLUSION

This work provides valuable insights into how underlying data, skewness, uncertainty, and visualization impact plea bargain decision-making. Firstly, building on prior work, we showed that participants adjusted their plea-bargain decisions based on the probability of conviction at trial when presented both with text and charts. Next, we found that participants adjusted their decisions based on the skewness of the distribution. Finally, we found that under time pressure, uncertainty in the form of text elicited the most risk-optimal decisions compared to other uncertainty visualizations, as well as

the mean only. We examined the impact of individual differences in numeracy, general risk propensity, and demographic factors, as well as the impact of time pressure on plea bargain decision-making. Our findings highlight that individual differences in numeracy and general risk propensity can impact decision-making under unlimited time conditions and time pressure. Altogether, our work broadens the scope of decision-support research and highlights the importance of considering alternative formats for data presentation.

ACKNOWLEDGMENTS

This work is supported by the Office of the Vice Chancellor at Washington University in St. Louis. This work was partially supported by the National Science Foundation (NSF) under Award No. 2142977 and 2330245, which funds the Engineering Research Center for Carbon Utilization Redesign through Biomanufacturing-Empowered Decarbonization (CURB).

REFERENCES

[1] [n. d.]. ABA 2023 Plea Bargain Task Force Report. https://www.americanbar.org/content/dam/aba/publications/criminaljustice/plea-bargain-tf-report.pdf. Accessed: 2010-09-30.

[2] David S. Abrams. 2011. Is Pleading Really a Bargain? *Journal of Empirical Legal Studies* 8, s1 (2011), 200–221. https://doi.org/10.1111/j.1740-1461.2011.01234.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1740-1461.2011.01234.x

[3] Nabil I Al-Najjar and Jonathan Weinstein. 2009. The ambiguity aversion literature: a critical assessment. *Economics & Philosophy* 25, 3 (2009), 249–284.

[4] Cynthia Alkon. 2016. Hard bargaining in plea bargaining: When do prosecutors cross the line. *Nev. LJ* 17 (2016), 401.

[5] American Bar Association. 2023. Plea Bargain Task Force Report. https://www.americanbar.org/content/dam/aba/publications/criminaljustice/plea-bargain-tf-report.pdf Retrieved from https://www.americanbar.org/content/dam/aba/publications/criminaljustice/plea-bargain-tf-report.pdf.

[6] Melanie Bancilhon, Zhengliang Liu, and Alvitta Ottley. 2020. Let's gamble: How a poor visualization can elicit risky behavior. In *2020 ieee visualization conference (vis)*. IEEE, 196–200.

[7] Melanie Bancilhon, Lace Padilla, and Alvitta Ottley. 2023. Improving evaluation using visualization decision-making models: A practical guide. In *Visualization Psychology*. Springer, 85–107.

[8] Melanie Bancilhon, Amanda Wright, Sunwoo Ha, R Jordan Crouser, and Alvitta Ottley. 2023. Why Combining Text and Visualization Could Improve Bayesian Reasoning: A Cognitive Load Perspective. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–15.

[9] Oren Bar-Gill and Omri Ben-Shahar. 2009. The Prisoners'(Plea Bargain) Dilemma. *Journal of Legal Analysis* 1, 2 (2009), 737–773.

[10] Oben K. Bayrak and John D. Hey. 2020. Decisions under risk: Dispersion and skewness. *Journal of Risk and Uncertainty* 61, 1 (2020), 1–24. https://doi.org/10.1007/s11166-020-09333-6

[11] Sarah Belia, Fiona Fidler, Jennifer Williams, and Geoff Cumming. 2005. Researchers misunderstand confidence intervals and standard error bars. *Psychological methods* 10, 4 (2005), 389.

[12] BELIV Workshop. 2024. BELIV Workshop 2024: Beyond Time and Errors – Novel Evaluation Methods for Visualization. https://beliv-workshop.github.io/about.html. Accessed: 2024-09-10.

[13] Stephanos Bibas. 2004. Plea bargaining outside the shadow of trial. *Harvard Law Review* (2004), 2463–2547.

[14] Susan AJ Birch and Paul Bloom. 2004. Understanding children's and adults' limitations in mental state reasoning. *Trends in cognitive sciences* 8, 6 (2004), 255–260.

[15] Tanja Blascheck, Lonni Besançon, Anastasia Bezerianos, Bongshin Lee, and Petra Isenberg. 2018. Glanceable visualization: Studies of data comparison performance on smartwatches. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 630–640.

[16] Kenneth S Bordens. 1984. The effects of likelihood of conviction, threatened punishment, and assumed role on mock plea bargaining decisions. *Basic and Applied Social Psychology* 5, 1 (1984), 59–74.

[17] Nadia Boukhelifa, Marc-Emmanuel Perrin, Samuel Huron, and James Eagan. 2017. How data workers cope with uncertainty: A task characterisation study. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 3645–3656.

[18] Adrian Bruhin, Helga Fehr-Duda, and Thomas Epper. 2010. Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica* 78, 4 (2010), 1375–1412.

[19] H Mitchell Caldwell. 2011. Coercive plea bargaining: The unrecognized scourge of the justice system. *Cath. UL Rev.* 61 (2011), 63.

[20] Lisa Cheong, Susanne Bleisch, Allison Kealy, Kevin Tolhurst, Tom Wilkening, and Matt Duckham. 2016. Evaluating the impact of visualization of wildfire hazard upon decision-making under uncertainty. *International Journal of Geographical Information Science* 30, 7 (2016), 1377–1404.

[21] Russell Covey. 2014. Behavioral Economics and Plea Bargaining. In *The Oxford Handbook of Behavioral Economics and the Law* (online edn ed.), Eyal Zamir and Doron Teichman (Eds.). Oxford Academic, Oxford. https://doi.org/10.1093/oxfordhb/9780199945474.013.0025

[22] Stephanie A Deitrick. 2007. Uncertainty visualization and decision making: Does visualizing uncertain information change decisions. In *Proceedings of the XXIII international cartographic conference*. 4–10.

[23] Evanthia Dimara and John Stasko. 2021. A critical reflection on visualization research: Where do decision making tasks hide? *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2021), 1128–1138.

[24] Michael Fernandes, Logan Walls, Sean Munson, Jessica Hullman, and Matthew Kay. 2018. Uncertainty displays using quantile dotplots or cdfs improve transit decision-making. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–12.

[25] Mirta Galesic, Rocio Garcia-Retamero, and Gerd Gigerenzer. 2009. Using icon arrays to communicate medical risks: overcoming low numeracy. *Health psychology* 28, 2 (2009), 210.

[26] Oren Gazal-Ayal and Avishalom Tor. 2012. The innocence effect. *Duke LJ* 62 (2012), 339.

[27] Anton Gollwitzer. 2018. The Role of Age in Plea Bargain Decision Making. *Available at SSRN 3149960* (2018).

[28] W Larry Gregory, John C Mowen, and Darwyn E Linder. 1978. Social psychology and plea bargaining: Applications, methodology, and theory. *Journal of Personality and Social Psychology* 36, 12 (1978), 1521.

[29] Marti A Hearst. 2023. Show It or Tell It? Text, Visualization, and Their Combination. *Commun. ACM* 66, 10 (2023), 68–75.

[30] Rebecca K Helm and Valerie F Reyna. 2017. Logical but incompetent plea decisions: A new approach to plea bargaining grounded in cognitive theory. *Psychology, Public Policy, and Law* 23, 3 (2017), 367.

[31] Jessica Hullman. 2019. Why authors don't visualize uncertainty. *IEEE transactions on visualization and computer graphics* 26, 1 (2019), 130–139.

[32] Jessica Hullman, Matthew Kay, Yea-Seul Kim, and Samana Shrestha. 2017. Imagining replications: Graphical prediction & discrete visualizations improve recall & estimation of effect uncertainty. *IEEE transactions on visualization and computer graphics* 24, 1 (2017), 446–456.

[33] Jessica Hullman, Paul Resnick, and Eytan Adar. 2015. Hypothetical outcome plots outperform error bars and violin plots for inferences about reliability of variable ordering. *PloS one* 10, 11 (2015), e0142444.

[34] Andrew Jordan, Ezra Karger, and Derek Neal. 2023. Heterogeneous impacts of sentencing decisions. (2023).

[35] Susan Joslyn and Jared LeClerc. 2013. Decisions with uncertainty: The glass half full. *Current directions in psychological science* 22, 4 (2013), 308–315.

[36] Susan Joslyn, Karla Pak, David Jones, John Pyles, and Earl Hunt. 2007. The effect of probabilistic information on threshold forecasts. *Weather and Forecasting* 22, 4 (2007), 804–812.

[37] Alex Kale, Matthew Kay, and Jessica Hullman. 2019. Decision-making under uncertainty in research synthesis: Designing for the garden of forking paths. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–14.

[38] Alex Kale, Matthew Kay, and Jessica Hullman. 2020. Visual reasoning strategies for effect size judgments and decisions. *IEEE transactions on visualization and computer graphics* 27, 2 (2020), 272–282.

[39] Alex Kale, Francis Nguyen, Matthew Kay, and Jessica Hullman. 2018. Hypothetical outcome plots help untrained observers judge trends in ambiguous data. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 892–902.

[40] Matthew Kay, Tara Kola, Jessica R Hullman, and Sean A Munson. 2016. When (ish) is my bus? user-centered visualizations of uncertainty in everyday, mobile predictive systems. In *Proceedings of the 2016 chi conference on human factors in computing systems*. 5092–5103.

[41] Martin G. Kocher and Matthias Sutter. 2006. Time is money—Time pressure, incentives, and the quality of decision-making. *Journal of Economic Behavior & Organization* 61, 3 (2006), 375–392. https://doi.org/10.1016/j.jebo.2004.11.013

[42] Michelle Korporaal, Ian T Ruginski, and Sara Irina Fabrikant. 2020. Effects of uncertainty visualization on map-based decision making under time pressure. *Frontiers in Computer Science* 2 (2020), 32.

[43] Besiki L. Kutateladze and Victoria Z. Lawson. 2018. Is a Plea Really a Bargain? An Analysis of Plea and Trial Dispositions in New York City. *Crime & Delinquency* 64, 7 (2018), 856–887. https://doi.org/10.1177/0011128717695224 arXiv:https://doi.org/10.1177/0011128717695224

[44] Isaac M Lipkus, Greg Samsa, and Barbara K Rimer. 2001. General performance on a numeracy scale among highly educated samples. *Medical decision making* 21, 1 (2001), 37–44.
[45] Mark J Machina and Marciano Siniscalchi. 2014. Ambiguity and ambiguity aversion. In *Handbook of the Economics of Risk and Uncertainty*. Vol. 1. Elsevier, 729–807.
[46] Hunter A McAllister and Norman J Bregman. 1986. Plea bargaining by defendants: A decision theory approach. *The Journal of social psychology* 126, 1 (1986), 105–110.
[47] Luana Micallef, Pierre Dragicevic, and Jean-Daniel Fekete. 2012. Assessing the effect of visualizations on bayesian reasoning through crowdsourcing. *IEEE transactions on visualization and computer graphics* 18, 12 (2012), 2536–2545.
[48] Melis Muradoglu, Joseph R Cimpian, and Andrei Cimpian. 2023. Mixed-effects models for cognitive development researchers. *Journal of Cognition and Development* 24, 3 (2023), 307–340.
[49] Limor Nadav-Greenberg and Susan L Joslyn. 2009. Uncertainty forecasts improve decision making among nonexperts. *Journal of Cognitive Engineering and Decision Making* 3, 3 (2009), 209–227.
[50] Alvitta Ottley, Aleksandra Kaszowska, R Jordan Crouser, and Evan M Peck. 2019. The curious case of combining text and visualization. *EuroVis 2019-Short Papers* (2019).
[51] Alvitta Ottley, Evan M Peck, Lane T Harrison, Daniel Afergan, Caroline Ziemkiewicz, Holly A Taylor, Paul KJ Han, and Remco Chang. 2015. Improving Bayesian reasoning: The effects of phrasing, visualization, and spatial ability. *IEEE transactions on visualization and computer graphics* 22, 1 (2015), 529–538.
[52] Lace M Padilla, Ian T Ruginski, and Sarah H Creem-Regehr. 2017. Effects of ensemble and summary displays on interpretations of geospatial uncertainty data. *Cognitive research: principles and implications* 2 (2017), 1–16.
[53] Anisha Patel. 2021. The effect of temporal discounting and loss aversion on mock plea bargain decision-making. (2021).
[54] Anthony Patt. 2001. Understanding uncertainty: forecasting seasonal climate for farmers in Zimbabwe. *Risk, Decision and Policy* 6, 2 (2001), 105–119.
[55] John W Payne, James R Bettman, and Mary Frances Luce. 1996. When time is money: Decision behavior under opportunity-cost time pressure. *Organizational behavior and human decision processes* 66, 2 (1996), 131–152.
[56] K. Petersen, A. D. Redlich, and R. J. Norris. 2022. Diverging from the shadows: explaining individual deviation from plea bargaining in the "shadow of the trial". *Journal of Experimental Criminology* 18 (2022), 321–342. https://doi.org/10.1007/s11292-020-09449-4
[57] Jonathan A. Rapping. 2012. Who's Guarding the Henhouse - How the American Prosecutor Came to Devour Those He Is Sworn to Protect. *Washburn Law Journal* 51, 3 (Summer 2012), 513–570.
[58] Valerie F Reyna and Charles J Brainerd. 1995. Fuzzy-trace theory: An interim synthesis. *Learning and individual Differences* 7, 1 (1995), 1–75.
[59] Maria Riveiro, Tove Helldin, Göran Falkman, and Mikael Lebram. 2014. Effects of visualizing uncertainty on decision-making in a target identification scenario. *Computers & graphics* 41 (2014), 84–98.
[60] Mark S Roulston, Gary E Bolton, Andrew N Kleit, and Addison L Sears-Collins. 2006. A laboratory study of the benefits of including uncertainty information in weather forecasts. *Weather and Forecasting* 21, 1 (2006), 116–122.
[61] Zick Rubin and Letitia Anne Peplau. 1975. Who believes in a just world? *Journal of social issues* 31, 3 (1975), 65–89.
[62] Ian T Ruginski, Alexander P Boone, Lace M Padilla, Le Liu, Nahal Heydari, Heidi S Kramer, Mary Hegarty, William B Thompson, Donald H House, and Sarah H Creem-Regehr. 2016. Non-expert interpretations of hurricane forecast uncertainty visualizations. *Spatial Cognition & Computation* 16, 2 (2016), 154–172.
[63] Lisa M Schwartz, Steven Woloshin, William C Black, and H Gilbert Welch. 1997. The role of numeracy in understanding the benefit of screening mammography. *Annals of internal medicine* 127, 11 (1997), 966–972.
[64] Winston Sieck and J Frank Yates. 1997. Exposition effects on decision making: Choice and confidence in choice. *Organizational Behavior and Human Decision Processes* 70, 3 (1997), 207–219.
[65] Meredith Skeels, Bongshin Lee, Greg Smith, and George Robertson. 2008. Revealing uncertainty for information visualization. In *Proceedings of the working conference on advanced visual interfaces*. 376–379.
[66] Megan L Small. 2022. The Effect of Magnitude and Probability on Plea Bargain Decision-Making. (2022).
[67] Laurence D Steinberg. 2014. *Age of opportunity: Lessons from the new science of adolescence*. Houghton Mifflin Harcourt.
[68] Chase Stokes and Marti A Hearst. 2024. Give text a chance: Advocating for equal consideration for language and visualization. *arXiv preprint arXiv:2404.00131* (2024).
[69] Chase Stokes, Vidya Setlur, Bridget Cogley, Arvind Satyanarayan, and Marti A Hearst. 2022. Striking a balance: reader takeaways and preferences when integrating text and charts. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2022), 1233–1243.
[70] René Levínský Tobias Brünner and Jianying Qiu. 2011. Preferences for skewness: evidence from a binary choice experiment. *The European Journal of Finance* 17, 7 (2011), 525–538. https://doi.org/10.1080/1351847X.2010.495478
[71] Ronald Wright and Marc Miller. 2002. The screening/bargaining tradeoff. *Stanford Law Review* (2002), 29–118.
[72] Don C Zhang, Scott Highhouse, and Christopher D Nye. 2019. Development and validation of the general risk propensity scale (GRiPS). *Journal of Behavioral Decision Making* 32, 2 (2019), 152–167.
[73] Tina M. Zottoli, Georgia M. Winters, Tarika Daftary-Kapur, and Conor Hogan. 2016. Plea Discounts, Time Pressures, and False-Guilty Pleas in Youth and Adults Who Pleaded Guilty to Felonies in New York City. *Psychology, Public Policy, and Law* 22, 3 (August 2016), 250–259.