This article is dedicated to the memory of the corresponding author's friend, Anat Kimchi, who tragically lost her life while the study was being conducted. Although Anat has never been a member of the research team, her research on probation conditions and obligations is a foundational piece of the current study.

Yan, S., Wilford, M. M., & Ferreira, P. A. (2022). Terms and conditions apply: The effect of probation length and obligation disclosure on true and false guilty pleas. *Journal of Experimental Criminology*. Advance online publication. https://doi.org/10.1007/s11292-022-09543-9

Terms and Conditions Apply: The Effect of Probation Length and Obligation Disclosure on True and False Guilty Pleas

Shi Yan[1], Miko M. Wilford[2], and Patricia A. Ferreira[3]

[1] School of Criminology and Criminal Justice, Arizona State University

[2] Department of Psychology, University of Massachusetts Lowell

[3] Department of Psychology, Iowa State University

Author Note

Shi Yan https://orcid.org/0000-0002-6484-7687

Miko M. Wilford https://orcid.org/0000-0002-8653-8893

Patricia A. Ferreira https://orcid.org/0000-0001-5511-0328

We have no conflict of interest to disclose. We thank Allison Redlich for her comments. A portion of these findings were presented at the 2022 annual meetings of the American Psychology-Law Society and the European Society of Criminology. We also thank Ariana Lowe, Hope Parker, and Olivia Shaw for their research assistance.

Correspondence concerning this article should be addressed to Shi Yan, School of Criminology and Criminal Justice, Arizona State University, 411 N Central Ave, Ste 600, Phoenix, AZ 85004, United States. Email: shiyan@asu.edu

Terms and Conditions Apply: The Effect of Probation Length and Obligation Disclosure on True and False Guilty Pleas

Abstract

Objectives: Although the literature suggests wrongful guilty pleas exist, less attention has been devoted to the false guilty pleas to probation sentences. We examined the plea decision-making process when participants faced probation.

Methods: We conducted a 2 (guilt status: innocent or guilty) x 2 (probation length: 1-year or 5-year) x 2 (probation obligations: general or detailed disclosure) between-participants experiment using an online sample (N = 906). Participants were randomly assigned to one of eight conditions and were asked whether they would accept a plea offer.

Results: Participants who were guilty and faced a shorter probation sentence were more likely to plead guilty. Participants receiving detailed disclosure were less likely to plead guilty, but the effect was less robust. Participants who faced longer sentence and received detailed disclosure reported larger declines in plea willingness.

Conclusion: The findings suggest that criminal defendants are sensitive to the length and disclosure of probation sentences.

Keywords: Guilty pleas; false guilty pleas; plea bargaining; probation; probation obligations; computer simulation

Introduction

The United States incarceration rate has become the centerpiece of criminal justice policy debates for decades, and the severity of incarceration is well-documented (Clear & Austin, 2017; Henson, 2020). However, the number of prisoners and jail inmates is overshadowed by the number of people serving community correction sentences. At any given time over the past two decades, there have been nearly twice as many Americans on probation as there have been incarcerated Americans (Kang-Brown et al., 2021; Oudekerk & Kaeble, 2021). Although most people view probation as a considerably more attractive alternative to imprisonment, people can face a list of potentially difficult conditions while serving on probation (Durnescu, 2011; Kohler-Hausmann, 2018; Wagner, 2010). Further, failure to abide by the terms of probation and parole can result in (re-)incarceration, as at least 153,000 individuals were incarcerated for non-criminal violations of their probation or parole in 2019 (Sawyer & Wagner, 2022). The term “mass probation” has started to emerge in the literature, calling for serious attention to the wide use of probation and its consequences (Phelps, 2017).

Regardless of sentence type, the overwhelming majority of criminal convictions result from a guilty plea rather than a jury or a bench trial (National Association of Criminal Defense Lawyers, 2018; Wilford & Bornstein, 2021). Supporters of plea bargaining see it as a tool to increase judicial efficiency and relieve defendants of harsher trial sentences (Church, 1979; Scott & Stuntz, 1992), while critics see plea bargaining as a coercive

practice that may drive even innocent defendants into pleading guilty (Hessick, 2021; McCoy, 2005; Rakoff, 2014). Despite disagreements concerning the overall prevalence of wrongful convictions and false guilty pleas (Cassell, 2018; Gould & Leo, 2010; "Kansas v. Marsh," 2006, J. Scalia's concurrence; Simon, 2012), researchers and practitioners have long recognized that false guilty pleas exist (Redlich, Bibas, et al., 2017). In fact, over a quarter of recorded exonerations (826 out of 3,250) involved false guilty pleas (National Registry of Exonerations, n.d.).[1] Although neither the number nor the proportion appears strikingly high, data on false guilty pleas, especially those not involving incarceration, are notoriously challenging to obtain (Gross, 2017; Wilford & Khairalla, 2019). Indeed, nearly every defendant pled guilty falsely and assisted by the Innocence Project served an extended prison sentence before exoneration (Innocence Project, 2009, n.d.). Due to the inherent challenges with the process, no researcher or policy organization has been able to provide a rigorous estimate of the number of people who served on probation after a false guilty plea.

Although we realize the difficulty of estimating the real-world number of people serving on probation due to a false guilty plea, we believe the number is not trivial for two reasons. The first is the sheer size of the probation population—nearly 3.5 million in 2019 (Oudekerk & Kaeble, 2021). Even if the false plea rate is the same for those incarcerated,

[1] Data updated on 10/3/2022 according to the National Registry of Exonerations website.

more people would be on probation while innocent, given the larger base rate. The second is the allure of probation as an incentive toward pleading guilty (Hessick, 2021; Roberts, 2011; Wilford & Khairalla, 2019), which further suggests that the false plea rate might be even higher when probation is offered as the sentence. The vast majority of studies on guilty pleas have focused primarily on carceral sentences (Bibas, 2004). Thus, the present study expands the scope of current plea research and provides a proof of concept that participant-defendants are responsive to the length and conditions of probation sentences.

Literature Review

Probation and False Guilty Pleas

Sociological, political, and legal scholars have frequently observed the incentives the state will provide in exchange for guilty pleas (Heumann, 1978; Nardulli et al., 1988)—these incentives are often referred to as “the plea discount” or “the trial penalty” (Yan & Bushway, 2018). Much of the scholarly work on plea bargaining recognizes one truism: the better the “deal,” the more likely the defendant is to plead guilty (Nardulli et al., 1988; Redlich, Bibas, et al., 2017; Redlich, Wilford, et al., 2017; Yan, 2020). Formally, the classic theoretical model, “bargaining in the shadow of trial” (often abbreviated as the shadow-of-trial or SoT model, Landes, 1971), predicts that the decision to plead guilty depends on

three variables: the sentence at trial (S_t), the sentence at plea (S_p), and the probability of conviction at trial (p). A rational defendant would plead guilty if

$$p > S_p/S_t$$

In plain language, more rational defendants would plead guilty if the plea deal is "attractive"—the difference between $S_p$and $S_t$being large enough so that taking the sentence S_p outweighs taking the risk at trial.[2] Many recent studies across disciplines (e.g., Abrams, 2011; Bonneau & McCannon, 2019; Bushway et al., 2014; Petersen et al., 2020; Wilford, Sutherland, et al., 2021; Yan, 2022) have been guided by the supposition that bargaining occurs in the shadow of the trial and provided empirical evidence supporting that rational defendants will generally be more willing to accept pleas that carry more lenient sentences. Unsurprisingly, qualitative and legal studies of plea bargaining have already documented many circumstances in which interested parties consider probation a less punitive option that incentivizes guilty pleas, suggesting its commonness (all emphases added):

> [An assistant prosecutor to the researcher:] I don't want to try these cases... So I tell the attorney, "Listen, if we try this case and your guy is convicted then I'll recommend jail. If you plead him, *I'll recommend probation*" (Church, 1976, p. 385).

[2] As the equation suggests, there are two ways to incentivize guilty pleas—reducing S_p or increasing S_t. Although the latter practice (often known as the trial penalty) bears significant legal and policy importance, it is beyond the scope of the current study and we will not discuss it here in depth.

> [A judge to the jurors:] “[L]adies and gentlemen, I want you to know, had he [the defendant] been willing to enter a plea on this case, *he would have been afforded probation*. He wouldn’t have gone to the institution. But now you’re going to the institution” ("State v. Scalf," 1998, pp. 622-623).

Notably, the prediction of the shadow model—“better” offers lead to more guilty pleas—applies to both guilty and innocent defendants (Wilford, Sutherland, et al., 2021). Both theoretical and empirical studies have found it impossible to design a plea bargaining system that perfectly separates innocent defendants from guilty ones (i.e., all guilty defendants plead guilty and all innocent defendants go to trial, Bjerk, 2007, 2021; Helm, 2022). When the size between the plea and trial sentences is sufficiently large, the *rational* decision, even for an innocent defendant, is to plead guilty to avoid the extra punitiveness at trial (Rakoff, 2014; Yan & Bushway, 2018). It is not rare for the prosecution to offer probation as a gesture of leniency in exchange for a guilty plea, especially for detained defendants (Henderson, 2019; Roberts, 2011). This strategy is often effective, even at times when the defendant is innocent. As Hessick (2021, p. 61, emphasis added) recounted from a public defender she interviewed,

> “As a public defender for over 7 years, I have yet to have a single client turn down a plea offer that gets them out of jail. *The chance to go home overwhelms the question of whether they are innocent or guilty*. Innocent people plead guilty all the time.”

Although the shadow model suggests that guilt status should not affect the rational defendant’s decision (except to the extent that guilt influences conviction probability), empirical research has shown that guilt status has a substantial impact on participant-

defendants' propensity to accept plea offers (Wilford, Sutherland, et al., 2021). In addition, studies have also shown that innocent defendants respond to many plea decision-making variables differently (e.g., attorney advice; Henderson & Levett, 2018; Henderson & Shteynberg, 2019; risk of COVID exposure in jail; Forston, 2022; Wilford, Zimmerman, et al., 2021; conviction probability; Wilford, Wells, et al. 2021). Innocent and guilty defendants may approach the decision to accept a plea offer in exchange for a non-carceral sentence differently.

The Conditions and Consequences of Probation

The conditions of a probation sentence are rarely as harsh as those within a jail or prison. Individuals on probation face less stringent rules than those housed in prisons and jails, and many people can (and are sometimes required to) reside in their own homes and maintain employment. However, this is not to say that probation is not onerous. A probation sentence can carry many terms and conditions, ranging from required educational and vocational training to drug and alcohol programs (Kimchi, 2019). Probation conditions can even be humiliating (Segelbaum, 2020) or costly (Harris et al., 2022; Ruhland et al., 2020). Moreover, people on probation are always at risk of potential revocation and subsequent incarceration (Kohler-Hausmann, 2018). Although many of those who have experience with the criminal justice system would overall prefer probation over incarceration (Petersilia & Deschenes, 1994; Spelman, 1995), some studies have found that they would *prefer a shorter incarceration sentence over a longer probation*

sentence due to the restrictions and conditions discussed above (Crouch, 1993; May & Wood, 2010). These findings demonstrate that receiving a probation sentence is a far cry from avoiding punishment (see also Hessick, 2021; Kohler-Hausmann, 2018; Yan, 2022).

A longstanding legal principle is that a guilty plea is valid only if the defendant makes the decision knowingly, intelligently, and voluntarily (Alschuler, 1975; "Boykin v. Alabama," 1969; Chin & Holmes, 2001). When prosecutors offer probation as an incentive for a guilty plea, many defendants may respond by accepting the offer after rational calculations. It is also possible that some defendants—including innocent ones—accept the offer without completely understanding what it means and entails. Redlich and Summers (2012) found that two-thirds of their sample, all of whom were actual criminal-justice-involved individuals, failed to achieve a 60% accuracy rate on their plea knowledge measure, and the *highest* score across the entire sample was 75%. Compared with carceral sentences, probation sentences have far more variation and less regulation (Kimchi, 2019). The substantial variation in probation conditions makes it even more difficult for defendants to understand and make knowing and intelligent plea decisions.

Given the scarcity of relevant administrative data concerning the prevalence of false guilty pleas in general, experiments are both common and appropriate for investigating factors driving false pleas (Redlich, Wilford, et al., 2017; Wilford & Khairalla, 2019). Yet, most experimental studies on guilty pleas have focused primarily on carceral sentences (e.g., Bushway et al., 2014; Garnier-Dykstra & Wilson, 2021; Wilford,

Sutherland, et al., 2021). Despite their ubiquity, there has been decidedly less attention devoted to the impact of probation sentences on plea decision-making. Moreover, studies on courts and guilty pleas have focused primarily on the length of the sentences, and fewer have explored their nature or conditions.

While no studies have experimentally examined the impact of disclosing probation terms on plea decision-making, two studies have tested the impact of disclosing collateral consequences on guilty pleas. Specifically, Edkins and Dervan (2018) found that disclosing a few potential collateral consequences had no significant impact on plea decision-making, particularly when pre-trial detention was a factor. Malone (2020), on the contrary, found that receiving information about additional collateral consequences of a criminal conviction—losing a nursing license, the right to vote, and eligibility for federal benefits—led to a significant decrease in guilty pleas. However, Malone's study did not mention pre-trial detention, and a carceral sentence would still be part of the plea offer. While both studies were inspiring and policy-relevant, neither focused specifically on the terms and conditions surrounding probation sentences. We address the gap as we detail below.

The Current Study

The current study builds on recent experiments (Edkins & Dervan, 2018; Malone, 2020) but furthers the literature in two directions. First, we examined the effect of

probation on plea decision-making by presenting a probation plea sentence to *all* participant-defendants in our study.[3] Second, we tested whether the effects of probation parameters (i.e., length and obligations) would differ between the innocent and the guilty. Unlike Edkins and Dervan (2018) and Malone (2020), we presented details regarding *the obligations of probation sentences* rather than *the collateral consequences of criminal convictions*. This is an important distinction criminal defendants may be so eager to avoid incarceration that the direct consequences of probation are given short shrift (as has been observed with collateral consequences of criminal convictions when defendants are in pre-trial detention; Edkins & Dervan, 2018). Past research has found that innocent participant-defendants are more prone to the influence of several plea-relevant variables than guilty defendants (Henderson & Levett, 2018; Wilford, Wells, et al., 2021; Wilford, Zimmerman, et al., 2021). Therefore, we hypothesized that,

1. Innocent participant-defendants would be more sensitive to probation length than guilty participant-defendants.
2. Innocent participant-defendants would be more sensitive to probation obligations than guilty participant-defendants.

[3] We are aware that there are a handful of other studies that have also presented a probation sentence to every participant-defendant (e.g., Helm et al., 2018; Zottoli & Daftary-Kapur, 2019). However, these studies were not focused on the impact of probation parameters on plea decision-making.

3. Guilty participant-defendants would be unaffected by probation length and obligations (i.e., they would be willing to accept plea offers at a high rate regardless of the probation manipulations).

We pre-registered all hypotheses and the main methodology (detailed below) on OSF (The Open Science Framework) prior to the start of data collection.[4]

Method

Pilot Study

To determine our manipulation conditions, we conducted a pilot study for an up-to-date estimate of the sentence exchange rate. The sentence exchange rate refers to how different sentence formats compare in value, similar to a currency exchange rate (e.g., five months of probation are equal to one month of incarceration). Although there has been a handful of exchange rate estimates published in the literature (Crouch, 1993; May & Wood, 2010; Spelman, 1995; Wodahl et al., 2013), these surveys were administered from the 1980s to the early 2010s. Thus, we surveyed 56 adults recruited on Prolific Academic and compensated them $0.67 for completing the pilot (which took approximately 5 minutes). Prolific Academic (www.prolific.co) is an online crowdsourcing platform commonly used in social science research (we used the same platform and target

[4] The pre-registration is available at https://osf.io/y8ps6/?view_only=0d7514cdbb19446b96189a8e20cef5d9

population for the main study). Participants were asked to provide the probation length equivalent (in harshness) to six-month and one-year jail sentences, as well as the jail length equivalent to one-year and two-year probation sentences. Then, from the four pairs of corresponding sentences (presented in Table 1), we estimated a power exchange function following the approach of Yan (2017).[5] With this power function, we were able to estimate the equivalent probation length for any given jail sentence length (or vice versa), and we present some examples in Table 1.

We found that, on average, participants perceived a one-year jail sentence roughly equivalent to 30 months (or two and a half years) on probation. Although all participants in the main study would ultimately be assigned a probation sentence in exchange for a guilty plea, we conducted the pilot study to ensure that the two potential probation sentences would be perceived differently relative to the trial sentence (which was fixed as a one-year jail sentence). We intentionally chose a probation sentence length that would be distinguishably less harsh (better) than the one-year jail sentence, and one that would be distinguishably harsher (worse) than this jail sentence. We aimed to include the latter

[5] A power function has the form of y = ax^b, where y is the severity of a given sentence, x is the length/amount, and a and b are parameters to be estimated. When x is zero, the value of y is also zero, indicating that a length or amount of zero would lead to zero in severity. When a is positive and b is between zero and one (our estimated a = 1.023 and b = 0.96), the function has a curvilinear upward shape that gradually flattens. Substantively, the severity increases with length/amount, but the increase in severity associated with each additional unit of length/amount gradually decreases (e.g., the additional month from 47 to 48 months is less harsh compared with the additional month from 3 to 4 months).

because studies have found that some defendants will choose any probation sentence—even one more punitive than the jail sentence—to avoid incarceration (Hessick, 2021). Based on the results of this pilot study, we chose to offer probation sentences of either 1 or 5 years.

Design and **Participants**

We employed a 2 (guilt status: innocent or guilty) x 2 (probation length: 1 year or 5 years) x 2 (probation obligations: general disclosure or detailed disclosure) between-subjects design. The disclosure manipulation refers to whether participants were provided general information regarding probationary sentences (e.g., various restrictions and requirements); or whether participants were provided more detailed information (e.g., restitution of $650). All participants were randomly assigned to one of the eight experimental conditions. They were compensated $4.75 for study participation (which required approximately 30 minutes total).

We determined the sample size via an a priori power analysis using G*Power 3.1 (Faul et al., 2009). The results showed that a sample size of at least 790 participants was needed to detect a small effect size for a two- or three-way interaction ($f^2 = 0.10$, $\alpha = .05$, $1 - \beta = .80$). We increased this estimate by 20% to account for potential loss of data due to incomplete responses, failed attention or manipulation checks, etc., resulting in a target sample size of about 950 participants.

We administered the study in Qualtrics (www.qualtrics.com), an online survey platform. Participants were required to be at least 18 years of age and to reside in the United States, with participants from the pilot being excluded from study participation. Participants first provided their informed consent and then responded to a set of pre-screening questions that assessed their eligibility to complete the study (i.e., age and country of residence). Eligible participants then provided their first name, which allowed participants to be referred to by their name at several points during the study to increase engagement. They were then directed to the simulation, and upon its completion, they were asked to respond to our dependent measures before being thanked and compensated.

In addition to the age and residency requirements, we embedded three attention checks throughout the study, each of which instructed participants to select a specific response (e.g., "If you are reading this question, select 'I prefer not to respond.'"). We also included three manipulation checks corresponding to our three manipulated variables. Participants were required to respond to at least two attention checks correctly and at least two manipulation checks to be retained in the final sample. A total of 97.1% of participants met these criteria, resulting in a final sample size of $N = 906$.[6] The sample had an average age of 35.1 years ($SD = 13.1$), with 53.6% identifying as male. The sample

[6] To check that our findings were not an artifact of our inclusion criteria, we also conducted the same set of analyses using only participants who passed all six attention and manipulation checks ($n = 704$). The main effect findings were similar to our full sample (see Supplemental File).

was predominantly White (70.3%), with Asian (9.5%), Black (7.2%), and multiracial (6.8%) as the next-largest categories. Over a quarter of the sample earned less than $20,000 per year (26.8%), and a combined fifth of the sample earned over $75,000 (21.7%).

Study Materials

An Interactive Simulation of Legal Procedures

The full legal scenario (i.e., the alleged crime, arraignment, and attorney meetings) was depicted via an animated computer simulation (Wilford, Sutherland, et al., 2021; Wilford, Zimmerman, et al., 2021). All the assets used for this simulation are freely available to researchers at researcher.pleajustice.org. Participants began by customizing an avatar that would represent them. Next, the simulation displayed participant-avatars walking toward a store in a shopping mall. While in the store, they asked the salesclerk for a pair of designer sunglasses in a locked display case. They then walked away to try on the sunglasses in front of a mirror. Shortly after that, they started receiving text messages stating they were late for a social engagement. They then looked back to see the salesclerk engaged with another customer.

After seeing a summons to appear in court, the simulation transitioned to a courtroom. In brief, the prosecutor claimed that the participant-avatar walked out of the store with a pair of designer sunglasses. The prosecutor stated that the sunglasses' value

warranted a felony-level larceny charge. The prosecutor further indicated that the salesclerk identified the participant-avatar. Video surveillance footage was then played, showing the participant-avatar walking toward the exit while wearing the missing sunglasses. Notably, the evidence was intentionally ambiguous: the salesclerk said that he gave the sunglasses to the participant-avatar (he did not say he saw the participant-avatar take them), and the surveillance footage showed the participant-avatar walking toward the exit with the sunglasses (but did not show them actually exit). When the arraignment concluded, participant-avatars experienced a flashback that revealed their actual guilt status. *Guilty* participants saw themselves walk out of the store wearing the designer sunglasses. *Innocent* participants, instead, saw themselves setting the sunglasses on a counter immediately before exiting the store.

The participant-avatars then met with a defense attorney. The defense attorney informed them that the prosecutor had extended a plea offer: they could either go to trial and risk being sentenced to one year in jail, or they could accept a plea deal, which would carry a probation sentence of *1 year* (short) or *5 years* (long). Participants were then asked to indicate the likelihood (on a scale of 0% to 100%) they would accept the plea offer. After doing so, they were provided a *general* or *detailed disclosure* of the probation terms and conditions (see below). Participants were then asked whether they wished to accept or reject the plea offer. Following their decision, they were asked to provide their willingness to accept the plea offer a second time.

Probation Terms and Conditions

The defense attorney informed all participant-defendants that probation sentences could generally involve various restrictions and requirements, such as monetary fees and periodic meetings; they were also informed that the judge may impose additional requirements or restrictions while they are on probation. All participant-defendants were also warned that if they violated probation conditions or were charged with any additional crimes, they could face further sanctions, including incarceration. This information made up the *general disclosure* of the probation terms and conditions.

To create a specific and realistic list of probation terms and conditions judges can assign to defendants, we searched extensively among several states' criminal statutes and their correctional department websites, the National Inventory of Collateral Consequences of Conviction (https://niccc.nationalreentryresourcecenter.org/), and recent news reports. We integrated the information from the sources and designed the following combination of probation obligations:

- One-time restitution of $650
- Monthly probation fee of $65 throughout the probation period (totaling $780 for the 12-month condition and $3,900 for the 60-month condition), over and above the restitution
- Meeting with the probation officer once every three months

- Maintenance of employment status for the entire probation period
- Firearm purchasing restrictions

The *detailed disclosure* of the probation obligations included all the items on this list in addition to the general disclosure.

Variables

The study focused on two sets of key dependent variables. The first is *plea acceptance*, coded dichotomously to capture whether each participant-defendant decided to reject (coded as 0) or accept (coded as 1) the plea offer at the conclusion of the simulation. The second set is the *willingness to accept the plea (WTAP)* measures. Twice during the simulation, we recorded participants' WTAP on a 100-point scale (i.e., "Given the plea deal you were offered and all the information from your defense attorney, please estimate the chances that you would plead guilty now as a percentage"; 0 = totally unwilling to plead, 100 = totally willing to plead). The first WTAP measure (WTAP1) was administered immediately after participants were offered the plea (i.e., after the plea sentence of 1 or 5 years of probation was offered). The second WTAP measure (WTAP2) was taken after participants were informed of the terms of the probation sentence (with either a general or detailed disclosure). The difference score, d(WTAP), was calculated by subtracting WTAP1 from WTAP2 for each participant (producing a possible range of -100 to 100). Importantly, this variable allowed us to capture the effect that disclosure of

probation terms had on participant-defendants' overall WTAP. A positive d(WTAP) value indicates that the participant-defendant became more willing to plead after receiving information regarding the probation terms. In contrast, a negative d(WTAP) value indicates that the participant-defendant became less willing to plead after receiving the details.

The key independent variables are the *three experimental factors*, as well as the three *two-way interaction effects* among them. To test for balance, we collected demographic and socioeconomic information from the participants, including *age* (coded continuously in years), *gender, race/ethnicity*, and *annual income level* (all coded categorically, see Table 2 for specifics).

Analytic Procedure

We began our analyses with general descriptive statistics (e.g., plea outcome frequencies) and tested for balance across our eight experimental conditions. We performed logistic (for plea acceptance) and linear (for WTAP-related outcomes) regressions to test for main effects of our experimental manipulations, as well as the predicted two-way interactions on plea outcome (dichotomous) and willingness to accept a plea (continuous).

Results

Balance and Raw Findings

Balance Among Groups

Our study design had eight between-participant conditions, which, if evenly assigned, should each include 113 participants. The actual group sizes varied between 98 and 130 participants. To test for successful randomization across conditions, we conducted a chi-squared test between the condition group and each of the categorical demographic variables (race/ethnicity, gender, income). We also conducted a one-way ANOVA on the continuous age variable among the groups, and present both sets of balance tests in Table 2. We did not detect a significant disparity among the groups on age, race/ethnicity, and income level at the .05 alpha level. However, we did detect that gender was slightly out of balance among the eight groups ($\chi^2(14) = 24.43$, $p = .04$), and that the disparity was consistently present regardless of our coding of gender (dichotomous or three-categorical). As a result, we included the gender variable (coded as female, male, and other) in all our subsequent data analyses.[7]

Raw Findings on the Outcomes

We present the study-wide descriptive statistics for the dependent variables in Table 2. Overall, 45.4% of the participant-defendants accepted the plea offer. The WTAP score at both time points (WTAP1 and WTAP2) hovered around the halfway point, with an

[7] We did not find a significant effect of gender in any of the models we estimated. Therefore, we omit further discussion from our manuscript to focus on the key findings of the experimental manipulations. In a series of sensitivity tests not presented in the manuscript, we found that omitting gender from the models led to no discernable differences in the main effects.

average score of 55.7% before the disclosure of probation terms and 48.6% after the disclosure. On average, participant-defendants decreased their WTAP by 7.1 percentage points after learning about the probation terms. It is also noteworthy that the actual range of d(WTAP) was -100 to +100, suggesting that some participant-defendants' WTAP changed completely after disclosure of the probation terms. Figure 1 presents the histograms of d(WTAP) by plea acceptance. Overall, most participant-defendants who accepted the plea had either an unchanged or increased WTAP2, and most participant-defendants who rejected the plea had a decreased WTAP2. Although the patterns are visibly different, it is also noteworthy that d(WTAP) is clearly not deterministic of the plea outcome. Some participant-defendants accepted the plea offer despite a decreased WTAP, and vice versa. The correlation between the two outcome variables is also moderate (r = .45), suggesting that they are distinct.

We further present the raw findings—the plea acceptance percentage and average d(WTAP) by experiment conditional—in Table 3. There were some visible differences in both dependent variables across the groups. The plea acceptance percentage ranged from 20.4% to 68.3%, while the average d(WTAP) ranged from -0.05 to -16.11. Among the eight conditions, the least favorable condition (innocent, 5-year, detailed disclosure) had both the lowest plea acceptance rate and the largest drop in d(WTAP), while the most favorable condition (guilty, 1-year, general disclosure) had the smallest drop in d(WTAP) and the second-highest plea acceptance rate (by 0.07 percentage point lower). Although

there appears to be some pattern, it is necessary to glean out the effect of each manipulation variable on the outcomes, which we present in detail below.

Effects on Plea Acceptance

Main Effects

Table 4 presents the main effects of the experimental conditions on the dichotomous plea acceptance outcome. We estimated two logistic regression models explaining plea acceptance, one containing only the main effects and the other containing the main effects and the three two-way interaction effects. For ease of interpretation, we present the marginal effect sizes calculated holding other regressors at their mean values alongside the coefficients. All three experimental manipulations had a statistically significant effect on plea acceptance. Guilty participant-defendants were 34.5 percentage points more likely to accept the plea offer than innocent participant-defendants (z = 9.64, p < .001, 95% CI [27.5, 41.5]; OR = 4.04, 95% CI [3.04, 5.38]). Participant-defendants facing the 5-year probation sentence were 16.4 percentage points less likely to plead than those facing the 1-year probation sentence (z = -4.60, p < .001, 95% CI [-23.4, -9.4]; OR = 0.51, 95% CI [0.39, 0.68]). Participant-defendants who received the detailed disclosure of probation obligations were 7.2 percentage points less likely to accept the plea than those who received the general disclosure (z = -2.01, p = .045, 95% CI [-14.2, -0.2]; OR = 0.75, 95% CI [0.56, 0.99]). Although the effect of

disclosure level emerged as significant, the magnitude of the effect was much smaller than the other two main effects.

Interaction Effects

We next added the three two-way interaction effects to the main model. The simple main effects of the 5-year probation sentence and detailed disclosure were negative, but their respective two-way interaction effects with guilt status were positive. Put together, this means that the respective effects of a longer probation sentence and detailed disclosure of probation were *smaller* for guilty participant-defendants than for innocent participant-defendants, as predicted. Meanwhile, the effect of detailed disclosure was larger for defendants who faced the 5-year probation sentence than for those who faced the 1-year sentence. However, all three interaction effects were small (marginal effect between 2.2 and 5.3 percentage points, *OR*s between 0.81 and 1.13) and non-significant ($p > .45$ for all three). In addition, none of the interaction effects were large enough to alter the sign of the corresponding main effects (e.g., $b_{5\text{-}year} = -0.63$ for innocent participant-defendants and $b_{5\text{-}year} = -0.41$ for guilty participant-defendants, both were still negative).[8] It is also noteworthy that the main effect of detailed disclosure was no longer significant. However, this change appears to be primarily due to an increase in the

[8] For both plea acceptance and d(WTAP), we also estimated a model including the three-way interaction effect among the three experimental condition variables. Neither was large or significant (nor was a three-way interaction predicted a priori), and the interpretation would be cumbersome. Therefore, we do not present these models in the manuscript.

standard error after the inclusion of the interaction effects (which increased by more than 70%) rather than a notable change in the main effect.[9] Thus, contrary to our primary hypotheses, guilt status did not significantly interact with either probation length or condition disclosure. While there was some evidence that innocent participants were more affected by the manipulations than guilty participants, guilty participants were still impacted by these manipulations.

Effects on Willingness to **Accept a Plea (WTAP) Variables**

Main Effects on d(WTAP)

Table 5 presents the findings on d(WTAP) for the three experimental conditions. The difference between guilty and innocent participant-defendants in d(WTAP) was not significant (t = 0.68, p = .49, 95% CI [-2.0, 4.2]). Participant-defendants who faced a 5-year probation sentence reported a 6.9-point larger decrease in WTAP than those facing a 1-year probation sentence. This difference was statistically significant (t = 4.29, p < .001, 95% CI [3.7, 10.0]). Participant-defendants who received detailed disclosure of probation terms dropped their WTAP by 7.7 percentage points relative to those who received a general disclosure. This difference was also statistically significant (t = 4.81, p < .001, 95% CI [4.6, 10.9]).

[9] In our supplemental analysis using Mize's (2019) approach, some main effects remained significant after the addition of interaction terms. See Supplemental File for details.

Interaction Effects on d(WTAP)

Table 5 also presents the findings on d(WTAP) after including the two-way interaction effects. Contrary to the findings on plea acceptance (and our hypotheses), the respective effects of the 5-year probation sentence and detailed disclosure were larger for guilty participant-defendants than for the innocent. However, both interaction effects were very small (b = -0.4 and -1.3, respectively) and non-significant ($p > .67$ for both). The effect of detailed disclosure on d(WTAP) was 4.9 percentage points larger for participant-defendants facing the 5-year probation sentence than those facing the 1-year sentence. Although this interaction was non-significant ($p = .13$), its size was larger than both corresponding main effects and much larger than the two other interaction effects. We also note that none of the main effects were significant in this model. However, it appeared again primarily due to the increased standard errors rather than major changes to the main effects. Thus, in contradiction to our primary hypotheses, guilt status did not significantly interact with either probation length or condition disclosure.

Exploratory Analyses on WTAP1 and WTAP2

To better capture the mechanisms underlying our findings on d(WTAP), we further examined the effects of our independent variables on WTAP1 (WTAP before the disclosure of probation obligations) and WTAP2 (WTAP after the disclosure of probation obligations) separately and present the findings in Table 5. Guilt status was the primary

determinant of WTAP at both time points, as guilty participant-defendants were much more likely to plead guilty than innocent ones (b_{guilty} = 23.6 for WTAP1 and b_{guilty} = 24.7 for WTAP2, both p < .001), and there appeared to be little change in the magnitude of the effect of guilt status before and after the probation obligation disclosure. However, the effects of probation length and probation obligation disclosure were both significant only after the disclosure (at WTAP2, $b_{5\text{-}year}$ = -9.6 and $b_{detailed\ disclosure}$ = -7.6, both p < .001). The null finding of detailed disclosure at WTAP1 was not particularly surprising, because none of the participants had received any information on probation terms and conditions when it was recorded. However, even though participants knew the probation sentence length they were to face at both time points, probation length only had an effect after the disclosure of probation obligations (i.e., at WTAP2). This finding supports the contention that laypeople generally do not understand the difficulties of probation; thus, they are only sensitive to changes in probation length after receiving information regarding the terms and conditions attached to probationary sentences.

Summary of Findings and Supplemental **Analyses**

Our analysis focused primarily on the role of three experimental manipulations (guilt status, probation length, and disclosure of probation terms and conditions) on two outcomes (plea acceptance and d(WTAP)). We found that all three experimental parameters were predictive of plea acceptance, but only probation length and probation obligation disclosure were predictive of d(WTAP), primarily through their effect on

WTAP2. Participant-defendants who faced the longer probation sentence and received more details of probation obligations were each less likely to plead guilty and reported a larger drop in WTAP after being informed of the conditions. Guilty participants were more likely to accept the plea than innocent ones, but they exhibited no larger or smaller change in WTAP. There did not appear to be significant interaction effects among the three variables on either of the dependent outcomes (contradicting our primary hypotheses).

To test the sensitivity of our findings, we repeated our analyses using the subsample of those who passed all six attention and manipulation checks (N = 704). Also, in response to recent calls against presenting and interpreting the marginal effects in logistic regression models on their face, we re-estimated the main and interaction effect sizes using the differences in predicted probabilities (Mize, 2019). We present both analyses in the Supplemental File appended to the manuscript.

Discussion

Wrongful convictions are costly. Using jury award and settlement data, a recent study estimated that the average cost of a wrongful conviction—which came with an average prison sentence of over 12 years—would be over six million dollars (Cohen, 2021). However, of the more than 200 wrongfully convicted individuals who were never incarcerated, *none* received any monetary compensation through litigation. While we

would never downplay the seriousness of wrongful incarcerations, we also believe the consequences of wrongful probations are more significant than the above data suggests. As many legal scholars have argued, the legal process is often arduous, and for many, the process itself is a punishment (Hessick, 2021; Kohler-Hausmann, 2018). The potential scope of wrongful probations is made more significant because most guilty plea decisions result in probationary sentences (Kang-Brown et al., 2021; O'Hear, 2008). The present study aimed to provide a direct link between wrongful conviction and mass probation. In addition to guilt status and sentence length—two relatively frequently-investigated variables in experimental plea research (Redlich, Wilford, et al., 2017)—we included a third manipulation of probation obligation disclosure.

The first noteworthy finding is that probation was a strong incentive to induce guilty pleas, regardless of the other experimental probation manipulations. Nearly a third of innocent participant-defendants pleaded guilty when doing so eliminated the risk of incarceration (by offering a probation sentence). Even among innocent participant-defendants who faced the five-year probation sentence—considerably harsher than the one-year jail trial sentence according to our pilot study—more than one-fifth accepted the plea offer just to avoid incarceration. We acknowledge the potential differences between our participant-defendants and actual defendants in their decision-making processes, and the possible differences between our numbers and the actual percentage of false guilty pleas to probation sentences. Nevertheless, the pattern resonates with reports from real-

world defendants regarding the overall incentivizing effect of probation on plea decision-making (Hessick, 2021; Roberts, 2011). Our study highlights the need to include probation sentences in the ongoing effort of deciphering and preventing false guilty pleas, given the willingness of defendants to take the plea that eliminates any risk of incarceration (regardless of guilt status).

Regarding the effects of the probation manipulations, participant-defendants who faced a longer probation sentence were less likely to accept the plea, which corresponds to a fundamental assumption of plea-bargaining—everything else being equal, an increase in plea sentence harshness will lead to a decrease in plea acceptance (Bushway et al., 2014; Redlich, Bibas, et al., 2017). However, the significance of this effect did not emerge until after the terms of probation were disclosed (i.e., at WTAP2). Furthermore, the detailed disclosure of the probation terms significantly lowered plea acceptance rates, though this disclosure had a larger impact on d(WTAP). These effects suggest that participant-defendants were receptive to their attorney's explanation of the conditions and terms they would have to face. While probation was perceived as a punishment, participants did not initially appreciate its significance. Only after receiving more information regarding the obligations associated with probation did participant-defendants fully appreciate the onerousness of an extended probation sentence (Kimchi, 2019; May & Wood, 2010; Wodahl et al., 2013; Yan, 2017).

We found guilty participant-defendants were more likely to plead than innocent ones, which is consistent with previous qualitative (Klein, 2006) and quantitative research (Bjerk, 2007; Wilford, Sutherland, et al., 2021; Wilford, Wells, et al., 2021). However, guilt status was unrelated to d(WTAP), suggesting that the introduction of additional information did not diminish its strong role in shaping WTAP. This finding is consistent with previous research positing that guilt status impacts the starting value (or reference point) from which defendants evaluate plea offers (Garnier-Dykstra & Wilson, 2021; Redlich, Bibas, et al., 2017). Moreover, participant-defendants who faced a longer probation sentence showed a bigger drop in WTAP after learning about probation terms, again suggesting that initially, laypeople might be unaware of how onerous probation can be (Redlich & Summers, 2012). These findings converge with previous research showing that increasing participants' awareness of the collateral consequences of a guilty plea reduced their likelihood of plea acceptance (Malone, 2020).

While we did observe the predicted effect of guilt status on plea outcomes, we did not find the predicted interactions. Specifically, we hypothesized that both probation conditions (i.e., length and obligations) would impact innocent participant-defendants more than guilty participant-defendants. Of note, guilty participant-defendants' rate of plea acceptance was lower than anticipated—a trend that could partially explain the absence of the predicted interaction. Even when participant-defendants were guilty and offered a short (1-year) probation term with general details of the demanding terms and

conditions, the plea rate was lower than 70%. This proportion is presumably much lower than the real-world rate, given the current domination of plea convictions (Wilford & Bornstein, 2021). Because the rate of true guilty pleas was relatively low (i.e., did not approach the ceiling), we were able to observe significant main effects of both probation manipulations on plea outcome and d(WTAP) for innocent and guilty participant-defendants. Overall, the results indicate that participant-defendants can be equally swayed by the perceived quality of the deal, regardless of guilt status. And, that those perceptions can be shaped by variations in even probation-level sentences.

Upon finding that two-thirds of their sample failed to achieve a 60% accuracy rate on their plea knowledge scale, Redlich and Summers (2012), stated:

> [A] significant proportion of defendants... may not fully understand and appreciate... the collateral consequences, the voluntariness associated with pleading guilty, and other important safeguards. As we see it, the problem is not so much with the standards set forth but rather is likely to lie in the *translation* and *interpretation of*, and *fidelity to*, the standards... Are [judges and attorneys] conveying the information in a manner that is understandable and meaningful to defendants? In turn, are defendants actively engaged in understanding, or does the resultant outcome (reduced charges, *deincarceration*) become the most important consideration above all else? In many ways, our findings generate more questions than answers (p. 640, all but one notes omitted, emphases added).

Our study answered some of the questions Redlich and Summers (2012) raised a full decade ago, bearing tremendous policy and intellectual significance. First, our study serves as yet another piece of evidence that to some innocent participant-defendants (Hessick, 2021; Rakoff, 2014), eliminating the risk of incarceration might indeed be “the

most important consideration above all else." Although wrongful probation is a lesser evil than wrongful incarceration, it is necessary for both researchers and policy advocates to begin formal evaluations regarding the damages and cost of wrongful probation sentences (Cohen, 2021). Second, participant-defendants' plea decisions can be affected by the number of sentence-related details disclosed. Currently, there is significant ambiguity regarding defense attorneys' obligations to disclose the terms of a probation sentence to their clients. Given the variation in probation obligations and conditions (Kimchi, 2019), it is crucial that courtroom actors—especially the defense and the judge—convey the specific list of conditions to each defendant clearly and completely. Third, the disconnect between plea acceptance and plea willingness in our sample suggests that some pleas might be *intelligent but not fully voluntary*. A dominant criticism of current plea-bargaining practices is how coercive the process can be (Rakoff, 2014; Yan & Bushway, 2018). Although providing a solution to this problem is beyond the scope of the present study, our findings could add to recent theoretical efforts on evaluating the key properties of valid guilty pleas—voluntariness, intelligence, and knowingness (Dezember et al., 2022; Luna, 2022).

This study is not without limitations. First, the study sample was not representative of a criminal-justice-involved sample, in which people of color and socioeconomically disadvantaged groups are typically overrepresented (United States Sentencing Commission, 2022). It is also possible that the sample generally did not perceive the

probation conditions we listed as particularly arduous, which could explain the relatively small effect of probation term disclosure on plea acceptance (cf. May & Wood, 2010). To criminal-justice-involved people, who are often financially-disadvantaged, our manipulations of monetary fees could have been more salient. Second, while the computer simulation we used is likely more immersive and engaging for participants than a vignette or a narrative (Wilford et al., 2019), it still poses a hypothetical scenario in which the consequences are not real. Thus, it is unclear whether participants were fully able to comprehend the magnitude of the decision with which they were faced. Third, for a myriad of reasons, it was difficult to determine how much information defendants typically receive regarding their probation terms prior to accepting a plea offer. It is also unclear at what pace this information might be conveyed and whether defendants truly comprehend all of the consequences of probation even after they are presented (Redlich & Summers, 2012).

To address the limitations, future studies would benefit from recruiting criminal-justice-involved and socioeconomically disadvantaged participants and validating the role of the key parameters of probation sentences. In addition, studies would also benefit from involving courtroom actor participants—defense attorneys, prosecutors, and judges. Both would help understand the general voluntariness, intelligence, and knowingness of plea decisions related to probation, and factors that may impede these properties of guilty pleas. Future research should also employ a variety of possible designs, including both

experiments (not limited to the animated simulation used in the present study) and systematic observations.

Ultimately, this work adds to a growing literature underscoring the importance of meaningfully *validating* guilty pleas. Defendants might nod their heads or utter the perfunctory one-word responses at a plea colloquy, but those responses do not mean that they are pleading knowingly and intelligently (Redlich & Summers, 2012). Research has already indicated that defendants can struggle to understand the general plea process (Redlich & Bonventre, 2015; Redlich & Summers, 2012; Zottoli & Daftary-Kapur, 2019), as well as the collateral consequences of pleading guilty (Edkins & Dervan, 2018; Malone, 2020). The results of this study highlight yet another area in which defendant comprehension must be meaningfully assessed. While probation offers an attractive "escape" from more adverse (proximal) consequences (i.e., incarceration), it also imposes numerous long-term (distal) conditions that, when unmet, can frequently result in incarceration (Sawyer & Wagner, 2022; Wilford & Wells, 2018). To preserve the integrity of our current justice system (termed a "system of pleas", *Lafler v. Cooper*, 2012), we must ensure that defendants accepting a plea offer for a non-carceral sentence understand and knowingly accept all the subsequent terms and conditions.

References

Abrams, D. S. (2011). Is pleading really a bargain? *Journal of Empirical Legal Studies*, *8*(s1), 200-221. https://doi.org/10.1111/j.1740-1461.2011.01234.x

Alschuler, A. W. (1975). The Supreme Court, the defense attorney, and the guilty plea. *University of Colorado Law Review*, *47*(1), 1-71.

Bibas, S. (2004). Plea bargaining outside the shadow of trial. *Harvard Law Review*, *117*(8), 2463-2547. https://doi.org/10.2307/4093404

Bjerk, D. (2007). Guilt shall not escape or innocence suffer? The limits of plea bargaining when defendant guilt is uncertain. *American Law and Economics Review*, *9*(2), 305-329. https://doi.org/10.1093/aler/ahm010

Bjerk, D. (2021). Socially optimal plea bargaining with costly trials and Bayesian juries. *Economic Inquiry*, *59*(1), 263-279. https://doi.org/10.1111/ecin.12922

Bonneau, D., & McCannon, B. C. (2019). Bargaining in the shadow of the trial? Deaths of law enforcement officials and the plea bargaining process. *Working paper*. https://doi.org/10.2139/ssrn.3457809

Boykin v. Alabama, 395 U.S. 238 (1969).

Bushway, S. D., Redlich, A. D., & Norris, R. J. (2014). An explicit test of plea bargaining in the “shadow of the trial”. *Criminology*, *52*(4), 723-754. https://doi.org/10.1111/1745-9125.12054

Cassell, P. G. (2018). Overstating American’s wrongful conviction rate: Reassessing the conventional wisdom about the prevalence of wrongful convictions. *Arizona Law Review*, *60*(4), 815-864.

Chin, G. J., & Holmes, R. W., Jr. (2001). Effective assistance of counsel and the consequences of guilty pleas. *Cornell Law Review*, *87*(3), 697-742.

Church, T. W., Jr. (1976). Plea bargains, concessions and the courts: Analysis of a quasi-experiment. *Law & Society Review*, *10*(3), 377-401. https://doi.org/10.2307/3053139

Church, T. W., Jr. (1979). In defense of "bargain justice" *Law & Society Review*, *13*(2), 509-525. https://doi.org/10.2307/3053266

Clear, T. R., & Austin, J. (2017). Mass incarceration. In E. Luna (Ed.), *Reforming Criminal Justice* (Vol. 1, pp. 55-75). The Academy for Justice.

Cohen, M. A. (2021). Pain, suffering, and jury awards: A study of the cost of wrongful convictions. *Criminology & Public Policy*, *20*(4), 691-727. https://doi.org/10.1111/1745-9133.12559

Crouch, B. M. (1993). Is incarceration really worse? Analysis of offenders' preferences for prison over probation. *Justice Quarterly*, *10*(1), 67-88. https://doi.org/10.1080/07418829300091711

Dezember, A., Luna, S., Woestehoff, S. A., Stoltz, M., Manley, M., Quas, J. A., & Redlich, A. D. (2022). Plea validity in circuit court: judicial colloquies in misdemeanor vs. felony charges. *Psychology, Crime & Law*, *28*(3), 268-288. https://doi.org/10.1080/1068316x.2021.1905813

Durnescu, I. (2011). Pains of probation: effective practice and human rights. *International Journal of Offender Therapy and Comparative Criminology*, *55*(4), 530-545. https://doi.org/10.1177/0306624X10369489

Edkins, V. A., & Dervan, L. E. (2018). Freedom now or a future later: Pitting the lasting implications of collateral consequences against pretrial detention in decisions to plead guilty. *Psychology, Public Policy, and Law*, *24*(2), 204-215. https://doi.org/10.1037/law0000159

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Garnier-Dykstra, L. M., & Wilson, T. (2021). Behavioral economics and framing effects in guilty pleas: A defendant decision making experiment. *Justice Quarterly*, *38*(2), 224-248. https://doi.org/10.1080/07418825.2019.1614208

Gould, J. B., & Leo, R. A. (2010). One hundred years later: Wrongful convictions after a century of research. *The Journal of Criminal Law and Criminology*, *100*(3), 825-868.

Gross, S. R. (2017). What we think, what we know and what we think we know about false convictions. *Ohio State Journal of Criminal Law*, *14*(2), 753-786.

Harris, A., Pattillo, M., & Sykes, B. L. (2022). Studying the system of monetary sanctions. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *8*(2), 1-33. https://doi.org/10.7758/rsf.2022.8.2.01

Helm, R. K. (2022). Cognition and incentives in plea decisions: Categorical differences in outcomes as the tipping point for innocent defendants. *Psychology, Public Policy, and Law*, *28*(3), 344-355. https://doi.org/10.1037/law0000321

Henderson, K. S., & Levett, L. M. (2018). Investigating predictors of true and false guilty pleas. *Law and Human Behavior*, *42*(5), 427-441. https://doi.org/10.1037/lhb0000297

Henderson, K. S., & Shteynberg, R. V. (2019). Plea decision-making: The influence of attorney expertise, trustworthiness, and recommendation. *Psychology, Crime & Law*, *26*(6), 527-551. https://doi.org/10.1080/1068316x.2019.1696801

Henderson, W. F. (2019). Probation detainers in Philadelphia: A due process dud? *Drexel Law Review*, *12*(1), 129-169.

Henson, A. (2020). Meet them where they are: The importance of contextual relevance in prison-based parenting programs. *The Prison Journal*, *100*(4), 468-487. https://doi.org/10.1177/0032885520939294

Hessick, C. B. (2021). *Punishment without trial: Why plea bargaining is a bad deal*. Abrams Press.

Heumann, M. (1978). *Plea bargaining: The experiences of prosecutors, judges, and defense attorneys*. The University of Chicago Press.

Innocence Project. (2009). *When the innocent plead guilty*. Retrieved June 13, 2022 from https://innocenceproject.org/when-the-innocent-plead-guilty/

Innocence Project. (n.d.). *Why do innocent people plead guilty to crimes they didn't commit?* Retrieved June 13, 2022 from https://guiltypleaproblem.org/

Kang-Brown, J., Montagnet, C., & Heiss, J. (2021). *People in jail and prison in Spring 2021*. Vera Institute of Justice.

Kansas v. Marsh, 548 U.S. 163 (2006).

Kimchi, A. (2019). Investigating the Assignment of Probation Conditions: Heterogeneity and the Role of Race and Ethnicity. *Journal of Quantitative Criminology*, *35*(4), 715-745. https://doi.org/10.1007/s10940-018-9400-2

Klein, R. (2006). Judicial misconduct in criminal cases: It's not just the counsel who may be ineffective and unprofessional. *Ohio State Journal of Criminal Law*, *4*(1), 195-222.

Kohler-Hausmann, I. (2018). *Misdemeanorland: Criminal courts and social control in an age of broken windows policing*. Princeton University Press.

Lafler v. Cooper, 566 U.S. 156 (2012).

Landes, W. M. (1971). An economic analysis of the courts. *The Journal of Law & Economics*, *14*(1), 61-107.

Luna, S. (2022). Defining coercion: An application in interrogation and plea negotiation contexts. *Psychology, Public Policy, and Law*, *28*(2), 240-254. https://doi.org/10.1037/law0000345

Malone, C. (2020). Plea bargaining and collateral consequences: An experimental analysis. *Vanderbilt Law Review*, *73*(4), 1161-1208.

May, D. C., & Wood, P. B. (2010). *Ranking correctional punishments: Views from offenders, practitioners, and the public*. Carolina Academic Press.

McCoy, C. (2005). Plea bargaining as coercion: The trial penalty and plea bargaining reform. *Criminal Law Quarterly*, *50*(2), 67-107.

Mize, T. (2019). Best practices for estimating, interpreting, and presenting nonlinear interaction effects. *Sociological Science*, *6*, 81-117. https://doi.org/10.15195/v6.a4

Nardulli, P. F., Eisenstein, J., & Flemming, R. B. (1988). *The tenor of justice: Criminal courts and the guilty plea process*. University of Illinois Press.

National Association of Criminal Defense Lawyers. (2018). *The trial penalty: The Sixth Amendment right to trial on the verge of extinction and how to save it*. N. A. o. C. D. Lawyers.

National Registry of Exonerations. (n.d.). *Exonerations in the United States Map*. Retrieved June 13, 2022 from

https://www.law.umich.edu/special/exoneration/Pages/Exonerations-in-the-United-States-Map.aspx

O'Hear, M. (2008). Plea bargaining and procedural justice. *Georgia Law Review*, *42*(2), 407-470.

Oudekerk, B., & Kaeble, D. (2021). *Probation and parole in the United States, 2019*. Bureau of Justice Statistics.

Petersen, K., Redlich, A. D., & Norris, R. J. (2020). Diverging from the shadows: Explaining individual deviation from plea bargaining in the "shadow of the trial". *Journal of Experimental Criminology*, *18*(2), 321-342. https://doi.org/10.1007/s11292-020-09449-4

Petersilia, J., & Deschenes, E. P. (1994). Perceptions of punishment: Inmates and staff rank the severity of prison versus intermediate sanctions. *The Prison Journal*, *74*(3), 306-328. https://doi.org/10.1177/0032855594074003003

Phelps, M. S. (2017). Mass probation: Toward a more robust theory of state variation in punishment. *Punishment & Society*, *19*(1), 53-73. https://doi.org/10.1177/1462474516649174

Rakoff, J. S. (2014, November 20). Why innocent people plead guilty. *The New York Review of Books*. http://www.nybooks.com/articles/2014/11/20/why-innocent-people-plead-guilty/

Redlich, A. D., Bibas, S., Edkins, V. A., & Madon, S. (2017). The psychology of defendant plea decision making. *American Psychologist*, *72*(4), 339-352. https://doi.org/10.1037/a0040436

Redlich, A. D., & Bonventre, C. L. (2015). Content and comprehensibility of juvenile and adult tender-of-plea forms: Implications for knowing, intelligent, and voluntary guilty pleas. *Law and Human Behavior*, *39*(2), 162-176. https://doi.org/10.1037/lhb0000118

Redlich, A. D., & Summers, A. (2012). Voluntary, knowing, and intelligent pleas: Understanding the plea inquiry. *Psychology, Public Policy, and Law*, *18*(4), 626-643. https://doi.org/10.1037/a0026066

Redlich, A. D., Wilford, M. M., & Bushway, S. (2017). Understanding guilty pleas through the lens of social science. *Psychology, Public Policy, and Law*, *23*(4), 458-471. https://doi.org/10.1037/law0000142

Roberts, J. (2011). Why misdemeanors matter: Defining effective advocacy in the lower criminal courts. *UC Davis Law Review*, *45*(2), 277-372.

Ruhland, E., Holmes, B., & Petkus, A. (2020). The role of fines and fees on probation outcomes. *Criminal Justice and Behavior*, *47*(10), 1244-1263. https://doi.org/10.1177/0093854820918877

Sawyer, W., & Wagner, P. (2022). *Mass incarceration: The whole pie 2022*. Prison Policy Initiative. https://www.prisonpolicy.org/reports/pie2022.html

Scott, R. E., & Stuntz, W. J. (1992). Plea bargaining as contract. *The Yale Law Journal*, *101*(8), 1909-1968. https://doi.org/10.2307/796952

Segelbaum, D. (2020, August 7). York County judge's sentence proposal: Hold sign that reads, 'I AM A SERIAL RETAIL THIEF.'. *York Daily Record*. https://www.ydr.com/story/news/watchdog/2020/08/07/york-county-judge-hands-down-unusual-sentence-repeat-shoplifter/3310771001/

Simon, D. (2012). *In doubt: The psychology of the criminal justice process*. Harvard University Press.

Spelman, W. (1995). The severity of intermediate sanctions. *Journal of Research in Crime and Delinquency*, *32*(2), 107-135. https://doi.org/10.1177/0022427895032002001

State v. Scalf, 126 Ohio App. 614 (1998).

United States Sentencing Commission. (2022). *2021 Annual report*. United States Sentencing Commission.

Wagner, A. K. (2010). The conflict over Bearden v Georgia in state courts: plea-bargaining probation terms and the specter of debtors' prison. *University of Chicago Legal Forum*, *2010*(1), 383-406.

Wilford, M. M., & Bornstein, B. H. (2021). The disappearing trial: How social scientists can help save the jury from extinction. *Psychology, Crime & Law*, *Advance online publication.* https://doi.org/10.1080/1068316x.2021.1984482

Wilford, M. M., & Khairalla, A. (2019). Innocence and plea bargaining. In V. A. Edkins & A. D. Redlich (Eds.), *A system of pleas: Social sciences contributions to the real legal system* (pp. 132-150). Oxford University Press. https://doi.org/10.1093/oso/9780190689247.003.0008

Wilford, M. M., Shestak, A., & Wells, G. L. (2019). Plea bargaining. In N. Brewer & A. B. Douglass (Eds.), *Psychological science and the law* (pp. 266-292). Guilford Press.

Wilford, M. M., Sutherland, K. T., Gonzales, J. E., & Rabinovich, M. (2021). Guilt status influences plea outcomes beyond the shadow-of-the-trial in an interactive simulation of legal procedures. *Law and Human Behavior*, *45*(4), 271-286. https://doi.org/10.1037/lhb0000450

Wilford, M. M., & Wells, G. L. (2018). Bluffed by the dealer: Distinguishing false pleas from false confessions. *Psychology, Public Policy, and Law*, *24*(2), 158-170. https://doi.org/10.1037/law0000165

Wilford, M. M., Wells, G. L., & Frazier, A. (2021). Plea-bargaining law: The impact of innocence, trial penalty, and conviction probability on plea outcomes. *American Journal of Criminal Justice*, *46*(3), 554-575. https://doi.org/10.1007/s12103-020-09564-y

Wilford, M. M., Zimmerman, D. M., Yan, S., & Sutherland, K. T. (2021). Innocence in the shadow of COVID-19: Plea decision making during a pandemic. *Journal of*

Experimental Psychology: Applied, *27*(4), 739-750. https://doi.org/10.1037/xap0000367

Wodahl, E. J., Ogle, R., Kadleck, C., & Gerow, K. (2013). Offender perceptions of graduated sanctions. *Crime & Delinquency*, *59*(8), 1185-1210. https://doi.org/10.1177/0011128709333725

Yan, S. (2017). Search for the hidden punishments: An alternative approach to studying alternative sanctions. *Journal of Quantitative Criminology*, *33*(1), 21-44. https://doi.org/10.1007/s10940-015-9275-4

Yan, S. (2020). Estimating the size of plea discounts: Why does it matter? In C. Spohn & P. K. Brennan (Eds.), *Handbook on sentencing policies and practices in the 21st century* (pp. 188-207). Routledge.

Yan, S. (2022). What exactly is the bargain? The sensitivity of plea discount estimates. *Justice Quarterly*, *39*(1), 152-173. https://doi.org/10.1080/07418825.2019.1707856

Yan, S., & Bushway, S. D. (2018). Plea discounts or trial penalties? Making sense of the trial-plea sentence disparities. *Justice Quarterly*, *35*(7), 1226-1249. https://doi.org/10.1080/07418825.2018.1552715

Zottoli, T. M., & Daftary-Kapur, T. (2019). Guilty pleas of youths and adults: Differences in legal knowledge and decision making. *Law and Human Behavior*, *43*(2), 166-179. https://doi.org/10.1037/lhb0000314

Figure 1. Histograms of d(WTAP) by Plea Acceptance

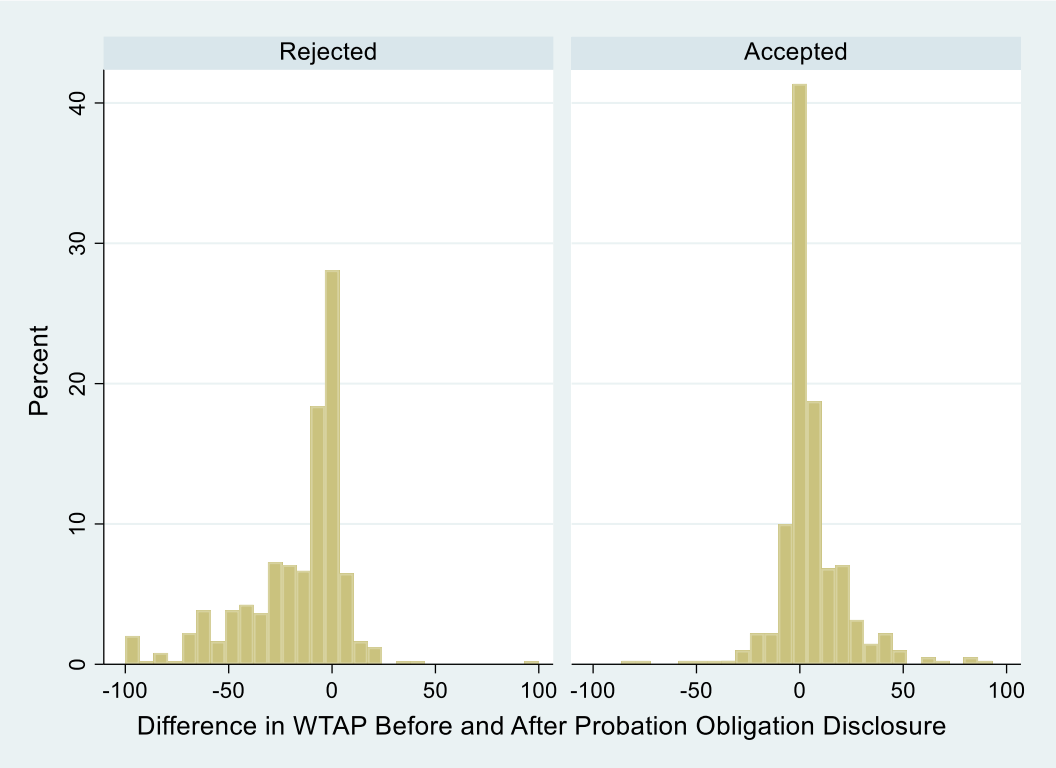

Table 1. Pilot Study Findings Regarding Probation and Jail Exchange Rate

	Mean
Observed from Pilot Study (*N* = 56)	
Months on probation equivalent to six months in jail	19.86
Months on probation equivalent to one year in jail	32.59
Months in jail equivalent to one year on probation	5.72
Months in jail equivalent to two years on probation	10.6
Predicted from the Power Function Model	
Months on probation equivalent to six months in jail	15.53
Months on probation equivalent to one year in jail	30.22
Months on probation equivalent to two years in jail	58.79

Note: Power function, following Yan (2017), predicted as

Probation Months = exp(1.023)*(Jail Months)^.96

Table 2. Descriptive Statistics and Balance

Variable	M	SD	Min	Max	Balance	Notes
Dependent Variables						
Plea Acceptance	0.454 [411]	0.498				Dichotomous, Accept = 1. For all categorical variables, number of participants in each category in brackets following the corresponding mean.
WTAP1	55.72	32.63	0	100		WTAP score before probation obligation disclosure
WTAP2	48.61	36.26	0	100		WTAP score after probation obligation disclosure
d(WTAP)	-7.11	24.59	-100	100		WTAP2 minus WTAP1
Demographic Variables						
Gender					$\chi^2(14) = 24.43, p = .04$	All balance tests are based on comparison across eight experimental conditions.
Female	0.435 [395]	0.496				
Male	0.536 [486]	0.499				
Other	0.028 [25]	0.164				
Age	35.13	13.11	18	83	$F(7, 898) = 0.95, p = .46$	
Race/Ethnicity					$\chi^2(56) = 61.81, p = .28$	

	0.703		
White	[637]	0.457	
Black	0.072 [65]	0.258	
Native American	0.003 [3]	0.057	
Asian	0.095 [86]	0.293	
Hawaiian/Pacific Islander	0.001 [1]	0.033	
Hispanic	0.043 [39]	0.203	
Other	0.008 [7]	0.088	
No response	0.007 [6]	0.081	
Multiracial	0.068 [62]	0.253	
Income			$\chi^2(35) = 46.66, p = .09$
	0.268		
Lower than $20,000	[243]	0.443	
$20,000 to $34,999	0.145 [131]	0.352	
	0.169		
$35,000 to $49,999	[153]	0.375	
	0.201		
$50,000 to $74,999	[182]	0.401	
$75,000 to $99,999	0.118 [107]	0.323	
Over $100,000	0.099 [90]	0.299	
N	906		

Table 3. Raw Results Across Experiment Conditions

Condition	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Guilt	Innocent	Innocent	Innocent	Innocent	Guilty	Guilty	Guilty	Guilty
Length	1 year	1 year	5 years	5 years	1 year	1 year	5 years	5 years
Disclosure	General	Detailed	General	Detailed	General	Detailed	General	Detailed
Plea Acceptance Percentage	40.77 [53]	31.63 [31]	23.53 [28]	20.37 [22]	68.22 [73]	68.29 [84]	61.11 [66]	47.79 [54]
Average d(WTAP)	-1.78	-6.64	-6.38	-16.11	-0.05	-6.06	-4.95	-15.73
n	130	98	119	108	107	123	108	113

Note: Numbers of participants in brackets next to corresponding percentages.

Table 4. Logistic Regression Models Explaining Plea Acceptance

	(1) Main Effects Only	(2) With Interactions
Guilty	1.40***	1.29***
	(0.15)	(0.24)
	[0.34]	[0.32]
Probation Length 5 Years	-0.66***	-0.63*
	(0.14)	(0.24)
	[-0.16]	[-0.15]
Detailed Disclosure	-0.29*	-0.23
	(0.14)	(0.25)
	[-0.07]	[-0.06]
Guilty * 5 Years		0.12
		(0.29)
		[0.03]
Guilty * Detailed Disclosure		0.09
		(0.29)
		[0.02]
5 Years * Detailed Disclosure		-0.22
		(0.29)
		[-0.05]
N	906	906

Note. Standard errors in parentheses, marginal effects in brackets, *** $p < .001$, ** $p < .01$, * $p < .05$

Table 5. OLS Regression Models Explaining WTAP Outcomes	(1) d(WTAP) Main Only	(2) d(WTAP) w Interactions	(3) WTAP1 Main Only	(4) WTAP2 Main Only
Guilty	1.10	1.87	23.58***	24.67***
	(1.60)	(2.74)	(2.03)	(2.24)
Probation Length 5 Years	-6.88***	-4.30	-2.71	-9.59***
	(1.61)	(2.70)	(2.03)	(2.25)
Detailed Disclosure	-7.72***	-4.62	0.14	-7.58***
	(1.61)	(2.79)	(2.03)	(2.25)
Guilty * 5 Years		-0.42		
		(3.21)		
Guilty * Detailed Disclosure		-1.33		
		(3.22)		
5 Years * Detailed Disclosure		-4.90		
		(3.21)		
N	906	906	906	906

Note. Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05