

**Race and Experimental Design: How Respondents may Read  
Context into a Neutrally Framed Scenario**

**Victor Volkman**

**April 18th, 2024**

**Abstract**

In the field of Law and Economics, experiments often use “neutral” language to avoid biasing respondents against certain actions such as stealing. However, a “neutral” frame for one group may not be the same “neutral” for another. In an online experimental study, we test the effect of framing on a two-player “stealing game” and compare the effect across samples of Black and White respondents. The behaviors of Black respondents in a neutral frame aligned much closer to those taken in a less serious “casino game” than in a more serious “police game” where actions were explicitly mentioned as “stealing”. This was in contrast to their White peers, who had distinct reactions to all frames. This suggests that past experiments may overstate the tastes for crime of Black respondents. How much a respondent trusted the police was also relevant. Ultimately, we caution that academics doing experimental work should consider whether experimental norms are overly tailored to certain groups. A better approach may be to embrace context.

**Keywords:** deterrence structure, priming, framing, criminology, risk aversion, communal norms, race, online, cooperation, policing, stealing game, lab experiment, perceived seriousness of wrongdoing, perceived seriousness of consequences, behavioral economics

**JEL Codes:** C91, D91, K10, K14, K42, J11, J15, and Z13

# 1 Introduction

External validity remains a massive concern for experimental works in the field of economics. Ideally, the design of an experiment should allow for preferences of respondents to be revealed through their behaviors. Yet, it is difficult to ensure all respondents will view a given decision the same way.

There is ample reason to believe context—the setting in which an experiment is conducted and how it is framed—affects the behavior of respondents (Meier, 2005). For example, respondents often behave differently when faced with decisions that can result in “losses” when compared to decisions where they may not “gain” as much. In several game types, respondents are more likely to shift away from strategies after experiencing the former rather than the latter, even if they result in the same ultimate payout. Research examining the experimental results of public good games has also suggested that framing a good as being part of an institution that a respondent is or was an active part of may make them more willing to contribute (Willinger and Ziegelmeyer, 1999; Ansink and Bouma, 2015).

Yet, considerations of context should not be limited solely to how an experiment is presented to respondents. Often the effect of framing can only be properly considered when its interaction with other key factors is also accounted for. For example, Alós-Ferrer et al. (2017) observed that framing outcomes in a game as “losses” only caused respondents to be more likely to shift away from strategies when they had the autonomy to choose their first strategy. Even when the explicit framing is the same, the implicit context of an economic behavior can differ.

Experiments like the one featured in Khadjavi (2014)—an experimental paper that observed the effect of imposing and later removing a system of punishment on the pro-social behavior of respondents—use “neutral” language and

avoided terms like “steal” and “caught stealing.” In theory, this avoids giving certain behaviors negative connotations, However, this approach assumes that the neutral scenario presented is similarly indicative of real world preferences for all groups. What if, instead, one group’s “neutral“ is not the same as another group’s? If one respondent interprets a decision in an experiment as a facsimile of the legal system but another interprets it as a game of chance or a sport, it is likely improper to pool all of the behaviors observed in that game as “stealing.“

Evidence suggests this type of heterogeneous evaluation may occur for key social groups. The framing of economic decisions has been shown to interact with gender in the past, even for seeming small changes such as pronoun choice (Chowdhury et al., 2017; Bruttel and Stolley, 2018; Bergh and Whichart, 2022). Additionally, life experiences often influence how scenarios are read (Bernile et al., 2017), and there is a wealth of evidence both for the life experiences of women being distinct from those of men (Matthews et al., 2000; Simon and Barrett, 2010) and that norms develop because of these experiences that have a tangible effect on economic behaviors (Everingham et al., 2007).

If this explanation holds, the experience of White and Black Americans may be even more disparate. Black youths and Black men in particular have very different talks with their parents about police when compared to their White peers (Sullivan et al., 2021), and Black families often put greater emphasis on preparing their children for interactions with officers—as noted in journals of Black psychology (Brunson and Weitzer, 2011; Harris and Amutah-Onukagha, 2019). The establishment of disparate communal norms across different racial groups has additionally been forwarded as a possible explanation for racial wealth disparities (Sowell, 1977) and the investment decisions of parents (Chiswick, 1983; Choi et al., 2002; Benjamin et al., 2010). It thus makes sense that the economic decisions of White and Black respondents observed in a lab setting may

be disparately informed, similar to the case with men and women.

However, there is a dearth of research that dissects the interactions between race and framing in this manner. A few papers have tested the claims of “self-categorization theory” experimentally by studying the effects of racial priming (Benjamin et al., 2010). Yet, the effect of gender is significant even in the absence of explicit primes (Chowdhury et al., 2017), whereas this largely hasn’t been considered with race. This is important because several papers show Black respondents to be more willing to go against moral conventions—including taking actions like stealing—compared to their White peers (Ones and Viswesvaran, 1998). These studies often use this “neutral” language. If Black respondents disproportionately read the scenario as being a game of chance or a sport, this result could be due to the tastes of crime of Black respondents being overstated.

To this end, we test the effect of framing on a two-player “stealing game” and compare the effect across samples of Black and White respondents. This game is a variant of the ultimatum game where Player 1 first proposes how to split an endowment. Player 2 responds by choosing to either accept the offer or attempt to “steal” from Player 1 at the risk of some monetary penalty. A similar variant has previously been used to measure the effects of imposed deterrence on pro-social behavior (Schildberg-Hörisch and Strassmair, 2012; Khadjavi, 2014).

Our treatment varies only the presentation of the monetary penalty and the action of “stealing”. In the more serious “police game”, stealing is explicitly called “stealing” and Player 2 has chance of being “caught.” In the less serious “casino game”, Player 2 is given a chance to “win chips” from Player 1 by spinning a roulette wheel. We compare stealing behaviors to a control group that uses the exact wording of Khadjavi (2014). We then look at the effects of treatment, the effects of race, and the interactions between race and treatment. All other aspects of the game including interactions between respondents and their

partners, the probability of detection, and the possible outcomes and payments for respondents are otherwise identical between these games.

Under the “police game” frame, we find that both Black and White respondents are significantly less willing to steal, but the effect of treatment is much more pronounced for Black respondents. We also find that Black respondents appear more likely to steal than their White peers, but this disappears when framing is taken into account. Taken as a whole, our results generally suggest that White respondents have distinct reactions to all three frames while Black respondents view the neutral frame as being similar to the less serious “casino game,” putting the racial differences in willingness to steal into perspective. This places the “neutrality” of our neutral frame into question and suggests that we may need to reconsider norms surrounding experimental procedure in the realm of law and economics.

## 2 Experimental Design

Our experiment leverages the design of Khadjavi (2014) in establishing a “deterrence structure” with a punishment mechanism to dissuade stealing. Specifically, deterrence is implemented such that the *probability* of detection is variable but *punishment* is constant.<sup>1</sup> Player 1 is the leading player, the Player 2 responds, and nature is a stochastic actor. At the start of the game, both players are given an endowment of  $E_1$  and  $E_2$  respectively, with  $E_1 > E_2$ . Player 1 then sends a signal containing their proposed division of funds  $W = \{w_1, w_2\}$ . If Player 2 accepts this offer, this will determine the ultimate payoffs of both players. Alternatively, Player 2 may reject this offer and attempt to take  $s$  from Player 1’s proposed endowment and transfer into her own. If this succeeds, Player 2’s final payoff will include this transfer and come out to  $w_2 + s$  while

---

<sup>1</sup>This is roughly equivalent to the measure  $Deter_{prob}$  from the 2014 paper.

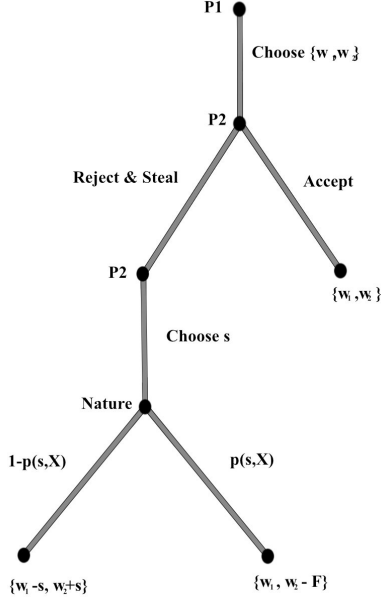


Figure 1: The Stealing game

Player 1 will be left with  $w_1 - s$  However, with probability  $p(s,X)$  this transfer will not work. In this case, Player 1 will receive their proposed endowment  $w_1$  as their final payoff, and Player 2 will incur a fee of  $F$  making their final payoff  $w_2 - F$ . Figure 3.1 illustrates the design of this game.  $P(s, X)$  is determined by the amount they have taken  $s$  from player 1's endowment and a normal random variable  $X \sim N(.25, .01)$  that will be uniquely drawn for each pairing.<sup>2</sup>

$$p(s, X) = \begin{cases} 0 & \text{if } s = 0 \\ X + 0.1s & \text{if } s > 0 \end{cases} . \quad (1)$$

<sup>2</sup>X value is unique to each respondent, and represents the baseline risk for an individual when engaging in “stealing”. Coupled with the use of the strategy method, a varying X value allows us to consider the marginal effect of a 1% increase in risk on trends in stealing behavior for each racial group at a given offer amount. This is meant to account for the possibility of different risk preferences between racial groups and possibly separate these differences from between group reactions to framing. The average for this normal random variable is based on the intercept of the probability function used in Khadjavi (2014).

This probability must be dependent on  $s$  to ensure that differences in the amounts taken by Player 2 are indicative of preferences and that Player 2 does not simply monotonically prefer to take the entirety of Player 1's endowment. The probability of being "caught" was presented for every amount the respondent chose to take in a numerical percentage. For example, a respondent in the control group with an  $X$  value of 0.25 and looking to take \$0.10 from Player 1 after receiving an offer of \$0 would see "This transfer has a 35% chance of success." The same player would see "You have a 35% chance of avoiding detection by the police." in the police game in the same scenario. In the casino game it would say "You have a 35% chance of winning the pot." No mention will be made of the normal random variable to either player.

In contrast to previous experiments, we conduct this game through the online survey platform Prodege. This was done to ensure the White American and Black American samples were drawn in a sufficiently wide and random manner. This also means we ask respondents to report how they would respond in strategy form. This entails having each respondent make 1 decision as Player 1 and decisions as Player 2 in 17 exhaustive cases given various offer amounts by their partner.<sup>3</sup> These cases were additionally presented in a random order to respondents in case of the effect of framing being biased by learning or other factors that may make the presentation order significant. Ultimate payoffs are based on these choices after respondents have been paired, and this was explicitly told to respondents to ensure they knew that their decisions had real weight.

The experiment begins with respondents completing a slider task. This determines each player's position in the experiment with the higher scoring

---

<sup>3</sup>Respondents were presented with 18 cases, but only 17 were used for this analysis. Regrettably, scenarios where the Player 1 offered \$1.30 needed to be omitted as there was a programming error that allowed respondents to attempt to steal more than was available to players exclusively in this scenario. This only affected the police game group but responses from all three groups were omitted. However, results in the \$1.30 scenario for the control group and casino game line up with reported trends.

responding acting as Player 1. This is standard procedure for stealing games as it means giving Player 1 a higher endowment and having them ultimately decide the opening split is not seen as arbitrary by the other player (Khadjavi, 2014).

This experiment was conducted on a sample of 292 White and 268 Black Americans in the General American population for a total sample size of 560<sup>4</sup>. Three treatments were randomly assigned to each sample, both differing from the control group in how the game is presented. In the control group, instructions given to respondents will mirror the “neutral“ language of Khadjavi (2014) closely. Specifically,  $s$  is described explicitly as a “transfer“ with a chance of “failing.“ In treatment 1—the “police game” scenario—Player 2 will be explicitly given the option to “steal” from Player 1.  $p(s,X)$  will be presented as the chance of being “caught by the police“. In treatment 2—the “casino game“ scenario—both players will be presented as playing a game in a casino. Player 1 will offer Player 2 some of her “chips“ and she can either take the offer or attempt to “win“ more of Player 1’s chips by spinning the roulette wheel. In all cases the fee  $F$  will be .20.

Table 3.1 details the mean values for various demographic categories between the White and Black samples. This largely confirms that samples are drawn in a manner that is similarly random.<sup>5</sup> However, there remain clear demographic differences in the samples, such as the higher proportion of women in the Black sample. This implies recovering the effect of treatment will likely require a deeper analysis than a simple comparison of means.

---

<sup>4</sup>The target for the sample size was 528, derived from the effect of racial priming on the risk preferences of Black respondents found in Benjamin et al. (2010) at a power of 90%, assuming that the true effect is 55% of the reported effect. It is additionally consistent with a 90% estimation of the effect reported in the replication of this study and a 65% estimation of the effect of deterrence found in Khadjavi (2014), each similarly at a power of 90%. The additional 42 observations are due to runoff, which is why the sample doesn’t evenly divide into 6.

<sup>5</sup>Table 4.5 in the Additional Tables section additionally shows no evidence of a geographical skew between either racial groups or treatment groups.



	White	Black	p-value of difference
Respondent Offer as Player 1	.6014 (.3780)	.5914 (.3840)	.7575
Female	.4966 (.5008)	.6530 (.4769)	.0018*
Education Level	3.455 (1.091)	3 (1.046)	.0000***
Parent's income	.5342 (1.703)	2.656 (1.544)	.0000***
Democrat	.5342 (.7613)	1.063 (.8254)	.0000***
Age Group	4.521 (1.458)	3.563 (1.614)	.0000***
X Value	.2500 (.0111)	.2501 (.0113)	.9352
Believed decisions effected payoff	1.301 (.7449)	1.354 (.7013)	.3865
Believed partner was a real person	.9247 (.8859)	1.254 (.8408)	.0000***
Cared what partner thought	1.045 (.9460)	1.269 (.9050)	.0040**
N	292	260	

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table 1: Summary Statistics and T-test of Group Demographics

## 2.1 Analysis Plan

Our analysis will attempt to recover the effect of the two framing treatments on 2 main outcomes: Player 2's willingness to steal—a binary variable that measures whether any amount was stolen—and the amount stolen by a respondent as a percentage of what they were able to steal. This analysis uses a pooled dataset of the observed stealing behaviors in 17 of the cases presented to respondents. Estimates for these effects will be recovered through a logistic regression in the first case and an OLS regression in the second, clustering errors around respondent ID in both cases. In these regressions, amounts stolen for various offers by Player 1 were included as measures of stealing behavior to show how preferences are reactive to the scenario.

Regressions were run on the Black and White samples separately and jointly with a dummy variable for race. The set of possible variables to include in the

regression was determined through a combination of backward and forward step-wise model selection methods via R. Variables will be consistent between the Black, White, and Total samples for regressions of each main outcome, with the set of variables chosen to first minimize the average value of Mallows’s  $C_p$  between the three samples then maximize the average adjusted R-squared value. The step-wise algorithm will be run on a data set containing all available variables from the experiment and an exhaustive list of interaction terms. Across all of these models, the “Police Game”, “Casino Game”, “Female”, “Respondent Offer as Player 1,” “X Value,” and “Risk Preference Question” variables will be included by default. By laying out this method beforehand, we ensure that the choice of variables is not arbitrary and should alleviate the need for correcting for multiple hypothesis testing.

### 3 Results

Comparing the means of the six race-treatment groups, as reported in table 3.2, reveals treatment to have a clear effect on behavior. Result 1 reports the effects of treatment.<sup>6</sup> The main three results can be summarized as follows.

**Result 1 (Treatment Effects):** *An increase in the perceived “seriousness” of the framing for the game appears to lower both the average willingness to steal and the average percent stolen for both samples separately and the total sample.*

Respondents in the police game were consistently the least likely to steal and they stole less on average overall. Meanwhile, respondents in the casino game were the most likely to steal and they stole more of what was available. Behaviors reported from the control group consistently lined up somewhere in

---

<sup>6</sup>Average percent stolen and willingness to steal are averages calculated across all amounts offered to respondents across the 17 scenarios presented to them. While willingness to steal is a binary variable, average willingness to steal is not. Thus, t-tests are still used for this variable rather than a Fisher effect test

	Black	White	Total	p-value of difference (Black vs. White)
<b>Panel A: Average Willingness to Steal</b>				
Control	.781 (.279)	.671 (.287)	.718 (.288)	.006**
Police Game	.625 (.337)	.503 (.341)	.572 (.344)	.007**
p-value of difference (Control vs. Police Game)	.001**	.000***	.000***	
Casino Game	.803 (.233)	.745 (.278)	.774 (.257)	.113
p-value of difference (Control vs. Casino Game)	.581	.029*	.047*	
Average Across Treatments	.386 (.298)	.319 (.319)	.687 (.312)	.000***
<b>Panel B: Average Percent Stolen</b>				
Control	.363 (.221)	.301 (.194)	.328 (.207)	.033*
Police Game	.317 (.233)	.225 (.203)	.274 (.222)	.002**
p-value of difference (Control vs. Police Game)	.180	.008**	.014*	
Casino Game	.433 (.233)	.414 (.212)	.427 (.222)	.388
p-value of difference (Control vs. Casino Game)	.052	.000***	.000***	
Average Across Treatments	.386 (.243)	.319 (.222)	.351 (.234)	.000***
N	260	292	560	

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table 2: T-test of Treatment on Amount Stolen/Willingness to Steal

the middle, but whether this behavior was more in line with the police game or casino game differed between the racial groups and which measure of stealing behavior was considered. The significance of these differences also varied. The effect of the police game on willingness to steal was most consistent, reporting a difference from the control group that is significant at the 1% level for each sample, and only the Black sample reporting a difference that is almost but not quite significant at the 0.1% level. Result 2 reports the effects of race.

**Result 2 (Race Effects):** *Black respondents appeared more willing to steal on average and steal larger (higher percentage) amounts in all groups. However, this difference was not statistically significant in the casino game.*

In terms of total dollar amounts taken, Black respondents stole \$0.62 more in the control group, \$0.92 more in the police game and \$0.19 more in the casino game when compared to their White peers. They were also willing to steal in roughly 2 more cases in the control and police game, and about 1 case more in the casino game.

The significance and size of the difference in stealing between races was dependent on framing. Whilst being consistent with Result 1, the p-value of difference between the Black and White sample is higher than .1 for both the average percent stolen and average willingness to steal in the casino game. In contrast, the p-value of this difference is consistently reported to be below .01 in the police game and consistently below .05 in the control group. This is even clearer in regards to average willingness to steal as not only are both significantly different at the 1% level, but the p-value of difference in the control group is even lower than for the police game. Additionally, framing showed a significant effect on the average percent stolen for White respondents but not Black respondents. For both observed measures of stealing, the effect of the casino game frame was also only significant for White respondents. This leads us to Result 3.

**Result 3 (Race-Treatment Interactions):** *Black respondents in the control group exhibit significantly different stealing behaviors in the police game, but not in the casino game. In contrast, White respondents appear to behave in distinct ways in each frame.*

We now check if these results remain consistent when accounting for variables such as observed risk preferences, amount offered by Player 1, and the geographic composition of each group.

### 3.1 Regression Estimates

A set of 23 covariates was chosen and added to the regression alongside the effects of each treatment frame and race. This covariate set includes gender, education level, parent's income, marital status, age, state-based fixed effect, the time elapsed during the survey, time of day the survey was conducted, answers to risk preference questions, religious affiliation, and if the respondent expressed concern about the survey such as it being confusing or tedious. For the sake of presentation, only the effects of framing, risk preferences, amount offered to the respondent in the scenario, and race are explicitly reported.

The reported coefficients were generally consistent with Result 1. Logistic regressions reported that the police game had a significant and consistently negative effect on willingness to steal for both samples. OLS regressions reported that the casino game had a significant positive effect on the percentage amount White respondents stole out of what was available to them, but the same was not true for Black respondents.

While other effects may not have been significant for all groups, the reported coefficients for the effects of framing on percent stolen and the effect of the casino game on willingness to steal suggest stealing was most discouraged in the police game and least discouraged in the casino game. While the sign of the effect of

the casino game on the willingness to steal for Black respondents conflicts with this, that result was not significant.

Result 2, however, was not corroborated. Whether a respondent is Black or White was not shown to have a significant effect on either percent stolen or willingness to steal. The coefficient for the Black dummy variable is even negative in panel A, but the sign of this effect may be irrelevant since said effect appears negligible.

While race did not affect stealing directly, the significance of framing did differ by race, lining up with Result 3. The police frame was shown to have a more pronounced significant effect on Black respondents' willingness to steal than White respondents. Black respondents were 80.8% less likely to steal in the police game compared to the control whereas White respondents were only 58% less likely to steal. The casino game frame did not affect the willingness to steal of White respondents or Black respondents. However, it did significantly affect the percent White respondents took when they did steal, while this was not true for Black respondents.

There were also a few other results of note. The amount offered by Player 1 had a significant effect on willingness to steal for both White and Black respondents. However, it only affected the percentage amount that White respondents stole from Player 1's endowment from what was available, and this effect was much less significant. Risk preferences were significant for both measures of stealing.<sup>7</sup> However, risk preferences appeared more significant in determining

---

<sup>7</sup>Three questions were used in the exit survey that were not tied to payoffs but displayed options with much starker differences in risk and possible payoffs, asking if respondents preferred "\$4 guaranteed or a %50 chance at X" with X being \$6, \$8, and \$10 respectively. "Risk seeking" is a measure of how many times a respondent chose the more risky option over the safe option and thus ranges from 0 to 3. X value was not shown to have a consistently significant effect on stealing decisions yet this measure was. Given (at least perceived) risk preferences did appear to factor into stealing decisions, it may be the case that there was not enough variation in X value to properly capture the effect risk. Thus, our ability to disentangle the differences in the effect of framing between groups from differences in risk preferences is limited.

willingness to steal for White respondents than for Black respondents.

## 4 Robustness

We now test to see if these results are robust to alternate specifications.

### 4.1 Considering Scenarios Independently

The regressions mentioned in Section 3 considered differences in willingness to steal and percent stolen both across respondents and across 17 different scenarios for the same respondent. We can alternatively compare behaviors in specific scenarios separately and find the differences in stealing behaviors across groups when Player 1 offers a given amount.

The results in three key scenarios are reported in Table 3.4: The minimum offer (\$0.00), the maximum offer, (\$1.70), and the lowest offer that gives the responding player a higher share than Player 1 (\$0.80). This allows us to consider how stealing behaviors may be affected by the perceived fairness of a scenario, as stealing may be seen as a more understandable response or even a just punishment for an unfair offer by Player 1. This is consistent with the amount offered being significant in determining willingness to steal as shown in Table 3.3.

Figures 3.2 and 3.3 additionally graph the offer from Player 1 against the average percent stolen and willingness to steal of respondents in the 6 race-treatment groups. This allows us to see visually how stealing behaviors change when the offer is more generous. Result 1 appeared to hold when Player 1's offer was high as respondents stole less and were less willing to steal in the police game while the opposite or no effect was seen in the casino game for White and Black respondents respectively. However, when Player 1's offer was lower, this relationship didn't hold as clearly and it was less likely to hold as the offer

	Black	White	Total
<u>Panel A: Willingness to Steal</u>			
Police Game	-0.808** (0.269)	-0.580* (0.262)	-0.588*** (0.178)
Casino Game	-0.191 (0.263)	0.183 (0.299)	0.180 (0.192)
Amount Offered	-0.534*** (0.107)	-1.102*** (0.123)	-0.778*** (0.077)
X Value ( $\times 100$ )	-0.059 (0.105)	-0.125 (0.114)	-0.060 (0.704)
Risk Seeking	0.250* (0.121)	0.393** (0.133)	0.308*** (0.082)
Black = 1			-0.003 (0.033)
<u>Panel B: Percent Stolen</u>			
Police Game	-0.054 (0.033)	-0.017 (0.031)	-0.046* (0.022)
Casino Game	0.042 (0.032)	0.091** (0.035)	0.063** (0.023)
Amount Offered	0.017 (0.015)	-0.036* (0.017)	-0.011 (0.011)
X Value ( $\times 100$ )	-0.002 (0.011)	-0.027* (0.013)	-0.001 (0.008)
Risk Seeking	0.046** (0.014)	0.050** (0.016)	0.044*** (0.010)
Black = 1			0.020 (0.025)
N	3910	3927	8058

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table 3: Stealing Behavior - Effect of Framing



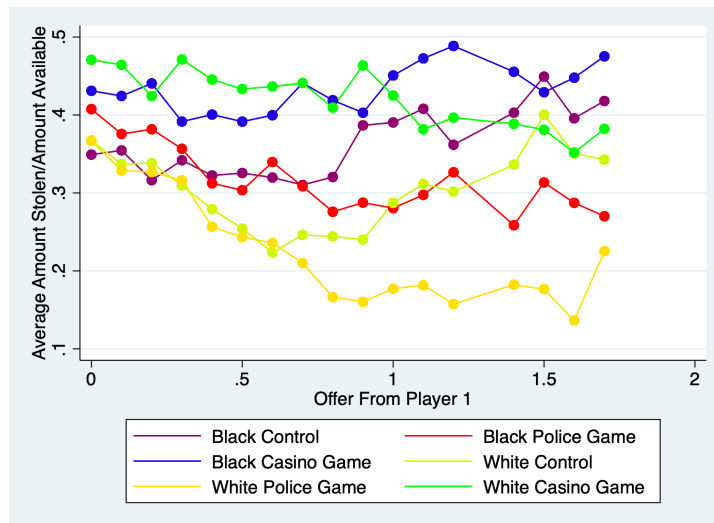


Figure 2: Effect of P1 Offer on Average Amount Stolen Over Amount Available

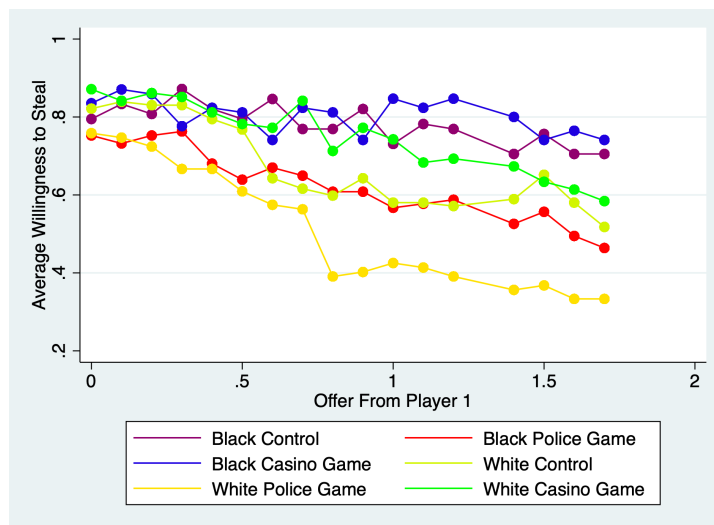


Figure 3: Effect of P1 Offer on Average Willingness to Steal

	<u>Amount offered:</u>					
	<u>\$0.00</u>		<u>\$0.80</u>		<u>\$1.70</u>	
	Black	White	Black	White	Black	White
<u>Panel A: Willingness to Steal</u>						
Police Game	0.580 (0.619)	-0.648 (0.689)	-1.605* (0.756)	-0.780 (0.489)	-1.266** (0.458)	-0.410 (0.437)
Casino Game	-0.212 (0.846)	1.577 (0.996)	-0.257 (0.724)	0.838 (0.498)	0.001 (0.546)	1.293** (0.443)
N	234	235	234	235	234	235
<u>Panel B: Percent Stolen</u>						
Police Game	0.085 (0.057)	0.020 (0.061)	-0.054 (0.059)	-0.013 (0.063)	-0.132 (0.068)	-0.034 (0.080)
Casino Game	0.032 (0.060)	0.098 (0.061)	0.028 (0.058)	0.186** (0.068)	0.034 (0.073)	0.125 (0.082)

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table 4: Stealing Behaviors Given Different Offers

decreased. Regarding willingness to steal, the effect of the police game frame was also much more significant for Black respondents than White respondents when Player 1's offer was high, supporting Result 3. Yet, the strength of this relationship decreased as the amount offered decreased in a similar manner. The significance of the effect of framing followed the same pattern. In contrast, the effect of the casino game frame on the percent stolen by White respondents was also corroborated but only for offers that were closer to an even split rather than for higher offers.<sup>8</sup>

## 4.2 Police Trust Index

Given the significance of framing appeared dependent on considerations of moral tastes like fairness, as shown above, respondents' opinions on the morality of police themselves were worth considering, especially given the very different

<sup>8</sup>Though only the results for these three scenarios are reported in this table, regressions were conducted for all of the 17 available cases. These confirm that the relationships noted here hold more when Player 1's offer increased in regards to willingness to steal and when Player 1's offer was closer to an even split in regards to percent stolen. However, the former was more consistent than the later. The casino game frame also had a more significant effect on the willingness to steal of White respondents when Player 1's offer was more than \$1.40

roles the institution has in the lives of White and Black individuals. Our results so far have suggested that Black respondents view the control game differently from White respondents, but is this due to differences in trust in police between races? To test this, we compare Black and White Respondents with high trust in the police and Black and White Respondents with low trust in the police.

The exit survey conducted after the experiment included questions relating to how respondents view police and how prominent a problem racial bias is for the police. The answers to these questions were compiled into a “Police Trust Index” with higher values relating to higher police trust<sup>9</sup>. In line with literature regarding the “police talk”, the average scores suggest White respondents to be more trusting of police than Black respondents across all treatment groups. The results reported in Table 3.5 show this difference to be generally highly significant.

Using these scores, we separated each racial group into those with high police trust and those with low police trust. Table 3.6 reports the effects of treatment on willingness to steal and the percentage stolen for each race-trust level group.<sup>10</sup> These generally line up with the findings in Table 3.3. Respondents generally stole less and were less willing to steal in more serious frames and race did not have much effect on stealing when accounting for framing. However, the effect of the police game frame was only significant for Black respondents with lower trust in the police while the police game frame was only significant for Black respondents with higher trust in the police.

---

<sup>9</sup>Police trust Index(PTI) is calculated as (In order of appearance on table 3.5) PTI = Question 1 + Question 2 - Question 3 - Question 4 - Question 5. PTI ranges from -15 to 15 points with each question ranging from -3 to 3 points with 3 representing “strongly agree”

<sup>10</sup>Again, these regressions include the same set of covariates as Table 3.3 but only explicitly report these effects for each panel. Here, logistic regressions would not converge when splitting up the sample based on police trust. Thus, the reported coefficient is from an OLS regression. The total sample additionally includes those who did not identify as White or Black

	Black	White	Total	p-value of difference (Black vs. White)
Police Trust Index	.273 (5.743)	5.511 (5.519)	2.956 (6.183)	.000***
“I trust the police”	.392 (1.955)	1.199 (1.648)	.820 (1.643)	.000***
“The police are a flawed but ultimately necessary part of society”	1.200 (1.678)	1.206 (1.609)	1.200 (1.633)	.999
“I am often afraid of the police”	-.012 (1.995)	-1.142 (1.851)	-.584 (2.000)	.000***
“I believe police treat me differently due to my race”	.904 (1.844)	-1.106 (1.968)	-.131 (2.150)	.000***
“I would probably make different life decisions if the police conducted themselves differently”	.427 (1.965)	-.858 (2.006)	-.222 (2.083)	.000***
N	260	292	560	

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table 5: T-test of Race on Police Trust Questions

	Black		White		Total	
	Low Trust	High Trust	Low Trust	High Trust	Low Trust	High Trust
<u>Willingness to Steal</u>						
Police Game	-0.102* (0.0498)	-0.058 (0.1039)	0.177 (0.0904)	-0.121* (0.0577)	-0.087* (0.0439)	-0.152 (0.0918)
Casino Game	0.053 (0.0506)	-0.060 (0.0916)	-0.093 (0.1427)	0.030 (0.0630)	0.014 (0.0445)	-0.006 (0.0467)
Amount Offered	-0.059** (0.0186)	-0.119*** (0.0314)	-0.129*** (0.0340)	-0.221*** (0.0272)	-0.082*** (0.0164)	-0.188*** (0.0209)
Black = 1					0.006 (0.0514)	0.050 (0.0639)
<u>Percent Stolen</u>						
Police Game	0.001 (0.039)	-0.043 (0.055)	0.108 (0.064)	-0.014 (0.035)	-0.040 (0.029)	-0.067* (0.032)
Casino Game	0.055 (0.040)	-0.003 (0.049)	0.004 (0.096)	0.102* (0.044)	0.015 (0.032)	0.076* (0.032)
Amount Offered	0.026 (0.017)	-0.003 (0.028)	-0.028 (0.028)	-0.040 (0.021)	0.006 (0.014)	-0.028 (0.017)
Black = 1					-0.030 (0.034)	0.079 (0.041)
N	2720	1258	1258	2737	4046	4046

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table 6: Effect of Framing - Low Police Trust vs. High Police Trust

## 5 Discussion

The results above generally point to the idea that Black respondents are seeing the casino game in our control, and not the “context-neutral” setting that White respondents may see it as.

The regressions in section 3 show framing to be very important but differences in race were negligible when accounting for framing. This puts the difference in the significance of framing for White and Black respondents into perspective. The police game frame had massive ramifications for the willingness to steal of Black respondents but the casino game affected neither willingness to steal nor the percentage stolen. White respondents, in contrast, exhibited different behaviors in all three frames, being less willing to steal in the police game and stealing a higher percentage of what’s available when in the casino game. This all points to the idea that Black respondents stole more not because this was reflective of a greater taste for crime or real life propensity to steal, but because they didn’t see a scenario that was as serious as their White peers did. When the seriousness of the situation was made clear, Black respondents changed their behavior accordingly.

While these results are robust to some specifications, framing may have been less important in cases where the responding player had less to lose by attempting to steal. Still, the fact that framing had an important interaction with moral tastes such as fairness is consistent with White and Black respondents having read very different contexts into the wording of the game in the control group. Black respondents stole less in the police game, but they stole even less when Player 1 had already given them a generous offer. This could be because that action felt “greedy” or “immoral” or it could be that the risk simply didn’t seem worth it. In any case, if this is where framing seemed most important, this is also where Black respondents appeared to see the game in our control as

similar to the casino game. In contrast, when Player 1's offer was sufficiently high White respondents saw the control and the police game fairly similarly but when the casino game was introduced they are significantly more willing to steal. Even if the tastes of crime of Black respondents weren't being overstated in all cases, there appeared to be a large split in the way the context of these experiments was being interpreted.

The significance of framing was also less pronounced for Black respondents with high trust in the police and White respondents with low trust in the police. While this may imply there might have been cases where Black and White respondents will act the same in the control scenario, it may also suggest the outcomes of the experiment might have been read differently based on race. For a Black respondent who doesn't trust the police, the risk of stealing may have been more pronounced because the risk of bodily harm felt higher. However, if White respondents only reacted to the police framing when they did trust the police, it might have been because they feared a loss of reputation rather than physical harm.

Overall, this overlooked interaction between race and framing suggests the willingness to steal of Black respondents may be overstated in previous experimental works. This should give us pause as to whether or not our experimental norms need to be re-examined, particularly concerning those using this "neutral" language. It is not clear that this is the same "middle ground" for all groups.

## **6 Conclusion**

Overall, the fact that reactions to framing were different for different racial groups should give researchers pause as to why we frame things the way we do. Neutral language is primarily used in experiments regarding Law and Economics

to guard against acquiesce bias, given the stigma associated with breaking the law. But do the benefits of this approach outweigh the costs associated with giving groups a scenario that they may read very differently? In this experiment not only did we observe that the actions of respondents in the control group arguably line up more with those of a game of chance than a facsimile of the legal system, but this appeared more true for Black respondents than White respondents. Furthermore, the factors that cause differences between Black and White respondents to be more pronounced are fairly logical given this understanding. In the police game, a Black respondent may weigh the consequences of being “caught” more heavily than a White respondent, and that means when the offer is high the benefit of stealing may not feel as worth the risk. Similarly, if your trust in police is low being “caught” breaking the law may have a more serious connotation, especially if the dynamic between police and race has been baked into your upbringing.

Oddly enough, while it makes intuitive sense that examples of “neutral” language are more tailored to White respondents than Black respondents, it is much easier to rationalize the effect of framing for Black respondents than White respondents. The significance of the effect of the police frame on willingness to steal for Black respondents increases logically with higher offer amounts and lower levels of police trust while patterns in the White sample are much more erratic. Yet, this does somewhat line up with expectations. Given the phenomenon of the “police talk,” it makes sense that as a group Black Americans may have a better thought out and more consistent understanding of which scenarios they can and can’t push their luck in, whereas White Americans range from those who have given it extensive thought and those who have never had to consider it. Still, this implies that results regarding tastes for crime from lab experiments using more “neutral” language may be overstated for Black

respondents.

The problem of framing is admittedly very difficult to address. In this experiment alone, we've shown that the effect of framing is very malleable to contextual factors such as the perceived fairness of said offer. Given the multifaceted nature of the effects of framing, we could be completely blind to several factors that might make a Black respondent read a different context into a scenario than their White peers. Asking for the perspective and raising the voices of our Black peers in academia may naturally lead to experiments feeling less tailored to respondents of a specific race, but Black and White respondents generally have very different life experiences and this will continue to inform their decisions and their expectations differently. The solution may instead be to drop the more sanitized "neutral" lab approach and embrace context. Though concerns over acquiescence bias are legitimate, and they often even differ across racial lines, we should strive to present these experiments to our respondents in the same way that we talk about them as academics.

## A Appendix



State	Total	Black	White	Control	Police Game	Casino Game
Florida	45	23	22	10	14	21
Illinois	33	18	15	11	13	9
Texas	42	24	18	16	13	13
New York	28	15	13	13	8	7
Georgia	37	24	13	12	13	12
Michigan	18	7	9	3	5	10
North Carolina	29	17	12	10	13	6
Ohio	25	14	11	7	6	12
Pennsylvania	28	11	17	12	8	8
Virginia	24	16	8	11	10	3
New Jersey	16	7	9	4	7	5
California	26	10	16	10	10	6
N	560	260	292	190	184	186

Table 7: Comparison of Geographic Distribution Between Groups

## References

- Alós-Ferrer, C., Hügelschäfer, S., and Li, J. (2017). Framing effects and the reinforcement heuristic. *Economics Letters*, 156:32–35.
- Ansink, E. and Bouma, J. (2015). Framed field experiments with heterogeneous frame connotation.
- Benjamin, D., Choi, J., and Strickland, A. (2010). Social identity and preferences. *The American economic review*, 100:1913–1928.
- Bergh, A. and Whichart, P. C. (2022). Mine or ours? unintended framing effects in dictator games. *Rationality and Society*, 34(1):78–95.
- Bernile, G., Bhagwat, V., and Rau, R. R. (2017). What doesn’t kill you will only make you more risk-loving: Early-life disasters and ceo behavior. *The Journal of Finance*, 72(1):167–206.
- Brunson, R. K. and Weitzer, R. (2011). Negotiating unwelcome police encounters: The intergenerational transmission of conduct norms. *Journal of Contemporary Ethnography*, 40(4):425–456.

- Bruttel, L. and Stolley, F. (2018). Gender differences in the response to decision power and responsibility—framing effects in a dictator game. *Games*, 9(2).
- Chiswick, B. (1983). The earnings and human capital of american jews. *Journal of Human Resources*, 18(3):313–336.
- Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2002). Defined contribution pensions: Plan rules, participant choices, and the path of least resistance. *Tax Policy and the Economy*, 16:67–113.
- Chowdhury, S., Jeon, J., and Saha, B. (2017). Gender differences in the giving and taking variants of the dictator game: Dictator giving, taking, and gender. *Southern Economic Journal*, 84.
- Everingham, C., Warner-Smith, P., and Byles, J. (2007). Transforming retirement: Re-thinking models of retirement to accommodate the experiences of women. *Women’s Studies International Forum*, 30(6):512–522.
- Harris, A. and Amutah-Onukagha, N. (2019). Under the radar: Strategies used by Black mothers to prepare their sons for potential police interactions. *Journal of Black Psychology*, 45(6-7):439–453.
- Khadjavi, M. (2014). On the interaction of deterrence and emotions. *Journal of Law, Economics, and Organization*, 31.
- Matthews, K., Raikkonen, K., Everson-Rose, S., Flory, J., Marco, C., Owens, J., and Lloyd, C. (2000). Do the daily experiences of healthy men and women vary according to occupational prestige and work strain? *Psychosomatic medicine*, 62.
- Meier, S. (2005). Does framing matter for conditional cooperation? evidence from a natural field experiment. *Contributions in Economic Analysis & Policy*, 5(2).

- Ones, D. and Viswesvaran, C. (1998). Gender, age, and race differences on overt integrity tests: Results across four large-scale job applicant data sets. *Journal of Applied Psychology*, 83(1):35–42.
- Schildberg-Hörisch, H. and Strassmair, C. (2012). An experimental test of the deterrence hypothesis. *Journal of Law Economics and Organization*, 28.
- Simon, R. and Barrett, A. (2010). Nonmarital romantic relationships and mental health in early adulthood does the association differ for women and men? *Journal of health and social behavior*, 51:168–82.
- Sowell, T. (1977). *Race and Economics*. Longman.
- Sullivan, J. N., Eberhardt, J. L., and Roberts, S. O. (2021). Conversations about race in Black and White US families: Before and after George Floyd’s death. *Proceedings of the National Academy of Science*, 118(38):1–4.
- Willinger, M. and Ziegelmeyer, A. (1999). Framing and cooperation in public good games: an experiment with an interior solution. *Economics Letters*, 65(3):323–328.