

**Punishing the privileged: Selfish offers from high-status allocators elicit greater  
punishment from third-party arbitrators**

Bradley D. Mattana<sup>a</sup>, Denise M. Barth<sup>b</sup>, Alexandra Thompson<sup>c</sup>, Oriel FeldmanHall<sup>d</sup>, Jasmin  
Cloutier<sup>b</sup>, & Jennifer T. Kubota<sup>b,e,1</sup>

**Affiliations:**

<sup>a</sup> Annenberg School for Communication, University of Pennsylvania, 3620 Walnut St.,  
Philadelphia, PA 19104, USA

<sup>b</sup> Department of Psychological and Brain Sciences, University of Delaware, 105 The Green,  
Newark, DE 19716, USA

<sup>c</sup> College of Liberal and Professional Studies, University of Pennsylvania, 3440 Market St., Suite  
100, Philadelphia, PA 19104, USA

<sup>d</sup> Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, 190  
Thayer Street, Providence, RI 02912, USA

<sup>e</sup> Department of Political Science and International Relations, University of Delaware, 18 Amstel  
Ave., Newark, DE 19716, USA

**1 Corresponding Author:**

Jennifer Kubota

Department of Psychological and Brain Sciences

Department of Political Science and International Relations

University of Delaware 105 The Green

Newark, DE 19716

Email: jennifer.kubota@gmail.com

Telephone: +1 (302) 831-4813

**Word Count:** 6,738

**Author Note:**

Data files and analysis scripts for all analyses including supplemental analyses are available on  
the Open Science Framework at <https://osf.io/9bxpk/>.

### **Abstract**

Individuals high in socioeconomic status (SES) are often viewed as valuable members of society. However, the appeal of high-SES people exists in tension with our aversion to inequity. Little experimental work has directly examined how people rectify inequitable distributions between two individuals varying in SES. The objective of the present study was to examine how disinterested third parties adjudicate inequity in the context of concrete financial allocations between a selfish allocator and a recipient who was the victim of the allocator's selfish offer. Specifically, this study focused on whether knowing the SES of the victim or the allocator affected the participant's decisions to punish the selfish allocator. In two experiments ( $N=999$ ), participants completed a modified third-party Ultimatum Game in which they arbitrated inequitable exchanges between an allocator and a recipient. Although participants generally preferred to redistribute inequitable exchanges without punishing players who made unfair allocations, we observed an increased preference for punitive solutions as offers became increasingly selfish. This tendency was especially pronounced when the victim was low in SES or when the perpetrator was high in SES, suggesting a tendency to favor the disadvantaged even among participants reporting high subjective SES. Finally, punitive responses were especially likely when the context emphasized the allocator's privileged status rather than the recipient's underprivileged status. These findings inform our understanding of how SES biases retributive justice even in non-judicial contexts that minimize the salience of punishment.

### **Keywords**

Socioeconomic Status, Justice, Decision Making, Inequality

## Introduction

The principle of equity is celebrated as a universal human value (1–3). Yet even in countries that prize equality, there are still observed preferences for some degree of inequity (4–10). This may be because the privileged have historically promoted the idea that hierarchy is established through non-exploitative means such as talent, hard work, or appointment from higher authority (11,12). However, when the privileged are exploitative (13) or even insufficiently generous (14,15), people often desire sanctions for these privileged perpetrators. For example, in 2018, the former extremely wealthy hedge fund manager and pharmaceutical CEO Martin Shkreli was convicted and sentenced to seven years in prison for defrauding hedge fund investors who trusted him with their money (16). Beyond Shkreli’s patently odious persona (17), prospective jury members seemed more concerned with Shkreli’s exploitation of the disadvantaged (viz., people who depend on affordable prescription drugs) rather than the privileged hedge fund investors he illegally defrauded (18). Shkreli’s public vilification for victimizing the disadvantaged highlights the growing public concern about the treatment of the “haves” and “have nots” (19,20) and raises the question of how the socioeconomic status (SES) of victims and perpetrators of inequity shapes the way third parties (e.g., judges, juries, court of public opinion) decide how to correct injustice. In particular, how does a perpetrator’s or victim’s SES shape punishment preferences when non-punitive solutions are also available (21)? To address these questions, the current study examined whether and how impartial third parties act to correct unfair distributions between individuals varying in SES. Based on previous work (21,22), we anticipated that participants would generally prefer non-punitive over punitive solutions, but that SES would nonetheless influence a third party’s willingness to punish. We predicted an increased preference for punitive solutions for highly unjust allocations if they were

perpetrated against the disadvantaged (i.e., low-SES individuals: Experiment 1) or committed by the privileged (i.e., high-SES individuals: Experiment 2). We also explored how the perceiver's SES impacted these preferences for punitive solutions to financial inequity.

Research shows that third parties are willing to punish perpetrators of inequity even when this might come at a personal cost (1–3,23–25). It has been argued that this sensitivity to fairness emerged relatively recently in our evolutionary history as human societies have grown in complexity (26), developing in tension with our heritage as a hierarchy-sensitive species (27). However, even non-human primates accustomed to relatively steep dominance hierarchies (e.g., macaques, chimpanzees) are sensitive to and take action to mitigate inequality (28). Given that social hierarchies are an intrinsic and systemic part of human culture (29) and organizations (30)—which help to shape our motives (30,31), affect (32,33), self-presentation strategies (34–36), evaluations of others (37–39), and support for redistributive socioeconomic policies (40–43)—it is critical to examine how inequity is perceived and addressed in social hierarchies. In the real world, systemic inequalities stemming from race and poverty (43–45) form the backdrop against which many fairness judgments are determined. Although several decades of research have examined fairness and punishment in the context of stratified social hierarchies (13,46–51), these studies typically focus on the participant's own status in that hierarchy (e.g., 47–49). Some research has focused on punishment decisions for people of varying social status levels (e.g., 13,46,50,51). However, these studies have tended to focus specifically on criminal offenders in an explicitly judicial context, showing that high-status perpetrators are seen as more intentional and/or callous, therefore deserving greater punishment than their lower status counterparts (13,46). This preference for punishing selfishness among the rich is further enhanced among those who generally desire greater equity (50).

Building on this literature, this is the first experimental investigation to our knowledge that directly manipulates the perceived status of perpetrators or victims in a context that minimizes the emphasis on criminality. This is important because such judicial contexts may predispose people to think in terms of punishment. Moreover, the degree to which the person's appearance or attributes reflect a particular criminal stereotype has been shown to influence punishment in judicial contexts (52,53). Accordingly, we tested whether and how participants acting as impartial third parties would rectify inequity in relatively low-stakes exchanges between partners varying in socioeconomic status (SES). Understanding whether and how inequity is punished in the absence of ascribed criminal behavior and when non-punitive means of correction are available (21,22) provides insight into baseline preferences for punishment for inequity on the basis of SES. Ultimately, this insight could then be applied to numerous real-world contexts involving sanctions for fairness violations. For example, SES and its proxies such as race (54–56) influence real-world sanctions across a wide range of contexts that vary in severity from bureaucratic inconvenience (57) to criminal sentencing (58–64). Thus, understanding how the victim's or perpetrator's SES contributes to fairness and punishment decisions is of paramount importance.

In two experiments ( $N=999$ ), participants on Amazon Mechanical Turk (MTurk) completed a third-party version of the Ultimatum Game, known as the Justice Game (21,22,65). Like the original Ultimatum Game, the Justice Game has been shown to capture the desire for punitive solutions to inequity (22,65). In the Justice Game, participants observed a series of single-shot exchanges between an allocator (Player A) and a recipient (Player B). Each exchange involved a split of funds between the players where Player A always determined how to split the funds with Player B. For example, Player A either acted quite selfishly, giving 20% to Player B

and keeping 80%, or only somewhat selfishly, giving 40% to Player B and keeping 60%. Participants played the role of an impartial third party (Player C) and had the power to respond to Player A's selfish offers in one of three ways: (a) punishment: Player A's allocation would be reversed such that Player B would end up with the largest share, and Player A would receive the smallest share; (b) compensation: the victim of the selfish offer (i.e., Player B) would receive additional funds to match the amount kept by Player A; and (c) accept: maintain the status quo and accept Player A's selfish offer. Participants were informed that the results of their decisions would be used to allocate bonuses to Players A and B, who ostensibly were real MTurk workers recruited prior to the experiment (see Supplemental Text S1). To make this cover story compelling, every trial involved a split of \$1 rather than \$10, which would have exceeded the participant's own compensation and prevailing payment rates on MTurk. Consistent with competing and paradoxical affinities for both fairness and hierarchical order, we posited that participants would (a) tolerate unequal distributions to a degree by preferring non-punitive redistributions, but (b) take strong punitive action against an allocator who perpetrated extreme selfishness (21).

To determine whether exploiting high-SES or low-SES victims is more likely to prompt punitive action, we manipulated the recipient's (Player B's but not A's) ascribed SES as either high or low (see Experiment 1 Methods for details). In line with public opinion about Martin Shkreli (18) and general perceptions of the rich as competent but cold (66,67), we anticipated diminished punishment of allocators who exploit high-SES (vs. low-SES) victims, especially when offers were extremely unfair.

To determine whether high-SES or low-SES perpetrators are more likely to prompt punitive action, we manipulated the allocator's (Player A's but not B's) ascribed SES as either

high or low (see Experiment 2 Methods for details). Research suggests that high-status individuals are expected to be more generous or at least non-exploitative toward low-status counterparts (15). When acting as perpetrators, high-SES perpetrators are seen as more motivated by selfish concerns than low SES perpetrators, which is associated with more severe punishments for rich offenders (46). Moreover, potentially due to status-based envy of the rich, evidence suggests that people find more pleasure when misfortunes (e.g., punishment) befall the rich versus the poor (68). We therefore predicted that participants would prefer to punish high-SES allocators, especially when offers were extremely unfair.

## Materials and Methods

For all experiments, we have reported all measures, conditions, data exclusions, and sample size determinations, consistent with best practices promoted by the Center for Open Science. Full details are provided in this section and in the online supplemental text. All analyses were conducted in R, version 3.5.3 (69) using the lme4 package for mixed-effects logistic regressions (70). Data files and analysis scripts for all analyses including supplemental analyses are available on the Open Science Framework at <https://osf.io/9bxpk/>. Furthermore, all research procedures complied with APA ethical standards and were approved by the Institutional Review Board at the University of Chicago, where the data were collected. All participants provided informed consent to participate in this research and agreed to their data being used for analysis following debriefing. All consent and agreements with research participants were obtained online using IRB-approved electronic consent and debriefing forms.

## Experiment 1

**Participants.** U.S.-based participants ( $n=502$ ) were recruited online using MTurk. These participants were required to have an 95% HIT approval rate and a minimum of 100 completed HITs. No other demographic or performance-based inclusion criteria were used for this study. MTurk samples tend to generate responses that are similar in reliability relative to data from lab-based samples (71–73). Additionally, MTurk provides a diverse participant pool in terms of age, ethnicity, and SES, something that is not always the case in American university participant pools (74). For Experiments 1 and 2, participants received a payment of \$2.00 USD after completing all parts of the experiment. Prior to analysis, we excluded data from three participants who either declined to authorize the use of their data for analysis or who failed to respond to the post-debrief data authorization question. Additionally, all trials with response latencies under 150 ms, resulting in the removal of 34 trials. Five participants with four or more of these latency-based exclusions were removed entirely from analysis, resulting in the removal of an additional 16 trials. These latency-based exclusions were implemented in both experiments to avoid the inclusion of physiologically implausible responses, responses made in the absence of deliberation, and the participants who made such responses for at least half of their trials. No other exclusion criteria were implemented. After these exclusions, the final sample consisted of 494 participants (265 Female, 228 Male, 1 Non-Binary;  $M_{age}=35.0$  years;  $SD_{age}=10.2$  years;  $Range_{age}=18–70$  years).

Our sample size was based on previous work by FeldmanHall and colleagues (21). Our final sample exceeded the largest participant group reported in their experiments. Using PANGAEA ([jakewestfall.org/pangea/](http://jakewestfall.org/pangea/)), we estimated that our 2 (SES: low, high)  $\times$  2 (Inequity: low, high) within-participants design (with one trial per condition) and final sample were



adequately powered to detect an effect as small as  $d=0.164$ ,  $1-\beta=.804$ , assuming a default variance parameter value,  $var(error) + var(participants*SES*inequity) = 0.417$ .

**Stimuli.** Eight White male faces from the Chicago Face Database (75) were chosen as stimuli for the modified Ultimatum Game described below (i.e., Justice Game). The eight faces were divided into two groups of four faces (i.e., one group per SES level) and equated for trustworthiness, attractiveness, and age based on pre-existing ratings from the Chicago Face Database. Trustworthiness and attractiveness ratings from the Chicago Face Database were provided on a 7-point scale from 1 (Not at all) to 7 (Extremely). Age ratings were provided as estimates (in years). The resulting two groups each consisted of four male faces that were somewhat below average in trustworthiness ( $M_1=3.17$ ,  $M_2=3.25$ ) and attractiveness ( $M_1=3.24$ ,  $M_2=2.86$ ), all two-tailed independent-samples  $|t(6)|<0.650$ ,  $p>.54$ . Faces from both groups were perceived as being approximately 27 years old ( $M_1=27.0$ ,  $M_2=27.2$ ), two-tailed independent-samples  $t(6)=-0.065$ ,  $p=.95$ . Faces belonging to the same group were superimposed on the same color background (i.e., blue or red) to indicate that face's SES level (see below). SES-color associations were counterbalanced across participants.

Previous research has relied on various antecedents of status including clothing, ascribed occupation/income/rank, body posture, facial structure, and even car ownership (39). Although such cues may afford ecological validity in some contexts, these status cues do not always unambiguously convey status level. Moreover, perceptual antecedents of status such as clothing are demonstrably confounded with important social dimensions like competence (76,77), making it difficult to reliably isolate effects of status. To avoid these potential pitfalls in our initial exploration of how status shapes justice decision making, we therefore used our existing procedure for ascribing status levels through learned color-coded background cues (78–82).

**Procedure.** Upon consenting to participate, participants first completed the SES–color association learning task (see below). Participants then read through some brief instructions describing the Justice Game and their role as Player C. Throughout the instructions, participants completed comprehension items that required the correct response in order to proceed (see Supplemental Text S1). To begin the main block of trials, participants were instructed to place their hands on the S, D, and F keys. After completing the Justice Game, participants completed several exploratory measures and demographic questions (see Supplemental Text S4). After completing these measures, participants were debriefed and given a completion code to receive compensation for completing the experiment.

**Learning status-color associations.** Prior to completing the Justice Game, participants first learned to associate the colors blue and red with different levels of SES (i.e., high or low). Both the SES–color association training and the subsequent Justice Game were presented online via Inquisit 4 Web (Version 4.0.9: Millisecond Software, Seattle, Washington). Participants initially read the following definitions of SES: “Those who have the HIGHEST social status tend to have the most money, the most education, and the most respected jobs. Those who have the LOWEST social status tend to have the least money, the least education, and the least respected jobs or no job.” Following these definitions, participants learned that they would see pictures of low- and high-SES individuals in the U.S., and that the pictures would be superimposed on a colored background denoting their SES level (e.g., blue=low SES, red=high SES). SES–color associations were counterbalanced across participants.

To thoroughly learn SES–color associations, participants completed simple association training blocks. In an initial block of 12 trials, participants passively viewed images of darkened silhouettes over a colored background (i.e., red or blue) paired with a sentence describing the

silhouette’s color-specific SES level (6 trials per SES level). Next, participants completed a block of 36 trials in which they viewed the same color-framed silhouettes without ascribed SES information (other than the colored background) and responded to a prompt regarding the silhouette’s SES level (e.g., “Does this color mean HIGH or LOW status in the US?”). Participants had unlimited time to press 1 for high SES or 2 for low SES. Incorrect responses elicited an error message: “INCORRECT—please give the correct response in order to proceed”. At the end of the block, participants received feedback on their overall accuracy and instructions that they would repeat the preceding block, irrespective of their initial accuracy. Any errors resulted in repetition of this training block. Training concluded with the next successful completion of the training block at 100% accuracy. On average, participants needed 1.3 blocks to reach 100% accuracy. Training data from two participants was unavailable. Although the present study did not test recall of status–color associations at the end of the experiment, previous research using this training procedure with relatively long cognitive tasks has shown good retention (~89%) and few differences when analyses were conducted with or without participants failing the post-task manipulation check (78).

**Third-party modified Ultimatum Game.** Based on the paradigm from FeldmanHall and colleagues (21), participants played the Justice Game as a third party with the power to redistribute resource allocations between two other individuals. On each of eight trials, participants decided whether and how to redistribute monetary splits that were proposed by an independent allocator (i.e., Player A). For each split, Player A divided \$1 USD between himself and a recipient (i.e., Player B). Participants were led to believe that Players A and B were real people and that Players A (i.e., a new Player A for every trial) made their allocations as part of a previous experiment (see Supplemental Text S1 for full task instructions). Moreover, each

participant was told that Players A and B would receive bonus payouts via MTurk as a direct result of the participant's decisions. In reality, allocator decisions were fixed. Player A's offers were always inequitable but varied in the degree to which they advantaged Player A versus Player B. Possible selfish offers were \$0.80/\$0.20 and \$0.60/\$0.40. Possible hyper-generous offers were \$0.20/\$0.80 and \$0.40/\$0.60. Although not the main focus of the present paper, we included hyper-generous offers (i.e., those favoring the recipient more than the allocator) because these offers allow the opportunity to examine inequity in the absence of a clear victim (see Supplemental Text S3 for analyses).

After viewing each monetary split, the participant (i.e., Player C) was then asked to determine the final monetary outcomes of Players A and B, choosing one of three options: accept, compensate, or reverse. The accept option maintains the status quo of Player A's original offer. In this case, the participant (i.e., Player C) accepts the monetary split between Players A and B that was offered by Player A. The compensate and reverse options signify different redistribution schemes. The compensate option bonuses whichever player received the least amount of money, increasing that player's earnings to match the amount received by the player initially advantaged by the allocator's (i.e., Player A) offer, thereby raising and equalizing the offer amounts between Players A and B. The reverse option flips the monetary split made by the allocator (i.e., Player A), effectively reversing the fortunes of Players A and B. Critically, the implications of the reverse option depend on the relative selfishness or generosity of the allocator's (Player A) offer. If the allocator (i.e., Player A) makes a selfish offer (e.g., \$0.80/\$0.20), then a reverse decision would punish Player A and reward the recipient of the selfish offer (i.e., Player B). Because our primary focus is on these selfish offers, we refer to the reverse option as the "punish" decision in the results section. If Player A makes a hyper-generous

offer (e.g., \$0.20/\$0.80), then a reverse decision would reward Player A but at the expense of Player B who was initially advantaged by the generous offer. For exploratory analyses and extended discussion of hyper-generous offers, see Supplemental Text S3.

Every trial began with a 250-ms fixation followed by the decision screen that remained until the participant responded. On the decision screen (see Fig 1), the allocator (i.e., Player A) was presented as a black silhouette of a face on a grey background. The recipient (i.e., Player B) was presented immediately to the right of Player A as a White male face framed by the color indicative of the face's SES level. Above the two faces was a consistent reminder to the participant that she/he was Player C. Immediately below the two faces, participants learned of Player A's initial allocation (e.g., "Player A decides to keep \$0.80 and offers \$0.20 to Player B.") Beneath this sentence, participants reviewed their response options, represented by three different keys (e.g., S=accept, D=reverse, and F=compensate). In the instructions, participants learn that the meaning of each key varies from trial to trial. To help avoid confusion regarding the changing meanings of the response keys, participants viewed the concrete consequences of each response key (e.g., "Player A gets \$0.80, Player B gets \$0.20") rather than its abstract representation (e.g., "accept"). The next trial began as soon as the participant made a response.

[Insert Fig 1]

**Fig 1. Example trial from the Justice Game from Experiment 1.** Here, the allocator (Player A) makes an extremely generous offer to the recipient (Player B). For exploratory analyses and extended discussion of hyper-generous offers, see Supplemental Text S3. The SES associated with the colored background for Player B was learned as part of the SES-color association training prior to the Justice Game. The colors representing high and low SES were counterbalanced across participants.

The trial sequence proceeded according to one of six randomly selected pseudorandom orderings intended to limit the repetition of factor levels (see Supplemental Text S1). Each

sequence presented an equal number of low- and high-SES recipients as Player B and each of the four possible allocations per SES level for a total of eight trials. The Justice Game concluded after the eighth trial.

**Data analysis.** Because the psychological meaning of redistribution decisions differs depending on whether the allocator's offer is selfish (i.e., allocating less to the recipient than to the allocator) or hyper-generous (i.e., allocating more to the recipient than to the allocator), we analyzed selfish and hyper-generous allocations separately. Analyses of decisions to punish selfish offers are reported in the main results section (see Supplemental Text S2 for analyses involving all other decisions for selfish offers), and exploratory analyses of hyper-generous offers are reported in Supplemental Text S3.

In separate mixed-effect logistic regressions, we tested how decisions to punish (punish = 1) versus not punish (compensate = 0, accept = 0) were predicted by offer inequity (extremely selfish = 1, moderately selfish = -1), recipient SES (high SES = 1, low SES = -1), and the SES  $\times$  Inequity interaction. All predictors varied within participants. The threshold for significance for Experiments 1–2 was set at  $p < .05$ . To the extent possible, we allowed for between-participants variance in intercepts and slopes for all within-participant factors (i.e., random effects) and the correlations between these random effects. However, sometimes, models failed to converge or were over-fitted (83). When a model failed to converge or was over-fitted, we followed the steps outlined in Supplemental Text S5. If the original maximal model converged and was not over-fitted, then we adopted that model in line with existing recommendations (84). The random effects structures for all reported models can be viewed in the study analysis script available on the Open Science Framework at <https://osf.io/9bxpk/>.

To investigate significant interactions, we conducted follow-up models on subsets of data corresponding to each level of each factor (Tables 1, 3, and 5) and each cell (Tables 2, 4, and 6) implicated in the interaction. As for the omnibus model, we attempted to fit as many random effects and correlation parameters as possible, following the aforementioned procedures. Because these follow-up models were restricted to decomposing significant interactions, we did not include any corrections for multiple comparisons.

## Experiment 2

**Participants.** U.S.-based participants ( $n=513$ ) were recruited online using MTurk. Eligibility criteria for Experiment 2 were the same as in Experiment 1 with the exception that participants from Experiment 1 were not eligible to participate in Experiment 2. Prior to analysis, we excluded data from one participant who declined to authorize the use of their data for analysis. Additionally, we excluded all trials with response latencies under 150 ms, resulting in the removal of 58 trials. Seven participants with four or more of these latency-based exclusions were removed entirely from analysis, resulting in the removal of an additional 18 trials. No other exclusion criteria were implemented. After these exclusions, the final sample consisted of 505 participants (273 Female, 230 Male, 2 Non-Binary;  $M_{age}=34.5$  years;  $SD_{age}=9.85$  years;  $Range_{age}=18-70$  years).

Our sample size was based on previous work by FeldmanHall and colleagues (21). As in Experiment 1, our final sample exceeded the largest participant group reported in their experiments. Using PANGAEA ([jakewestfall.org/pangea/](http://jakewestfall.org/pangea/)), we estimated that our  $2$  (SES: low, high)  $\times 2$  (Inequity: low, high) within-participants design (with one trial per condition) and final sample were adequately powered to detect an effect as small as  $d=0.162$ ,  $1-\beta=.803$ , assuming a default variance parameter value,  $var(error) + var(participants*SES*inequity) = 0.417$ .

**Design and analysis.** The face stimuli, status-color association training, Justice Game paradigm, and data analysis were the same as in Experiment 1. The only difference is that the recipient (i.e., Player B) was represented by a black silhouette and the allocator (i.e., Player A) was represented by the same face stimuli used in Experiment 1 (see Fig 2). In other words, Experiment 2 varied the SES level of the allocator rather than the recipient. Accordingly, the SES factor in all analyses for Experiment 2 refers to the SES level of the allocator rather than the recipient.

[Insert Fig 2]

**Fig 2. Example trial from the Justice Game from Experiment 2.** Here, the allocator (Player A) makes a somewhat selfish offer to the recipient (Player B). The SES associated with the colored background for Player A was learned as part of the SES–color association training prior to the Justice Game. The colors representing high and low SES were counterbalanced across participants.

To facilitate comparisons across experiments, we combined the data from Experiments 1–2 and repeated the same omnibus model used for these experiments but with experiment as an additional between-participants predictor (experiment 2 = 1, experiment 1 = -1). The resulting model included predictors for all possible interactions between SES, offer inequity, and experiment. As for the preceding analyses, we allowed for between-participants variance in intercepts and slopes for all within-participant factors (i.e., random effects) and the correlations between these random effects. However, sometimes, models failed to converge or were over-fitted (83). When a model failed to converge or was over-fitted, we followed the steps outlined in Supplemental Text S5. If the original maximal model converged and was not over-fitted, then we adopted that model in line with existing recommendations (84). The random effects structures for all reported models can be viewed in the study analysis script available on the Open Science Framework at <https://osf.io/9bxpk/>.



## Results

In the following experiments, we differentiated between the competing accounts of how SES shapes decisions to address inequity in financial offers from Player A to Player B from the perspective of a third party (Player C) who has no stake in the exchange—the court of public opinion. Although participants generally preferred non-punitive compensation over punishment (21,22,65), we nonetheless observed sensitivity to SES and offer inequity when participants did opt to punish selfish allocators. Accordingly, our main analyses focused on how frequently participants chose to punish selfish offers as a function of the offer’s selfishness and the SES of the recipient (Experiment 1) or the allocator (Experiment 2). Using mixed-effects logistic regression, we explored how SES and inequity influenced decisions to punish.

### Experiment 1: Effects of Recipient SES on Third-Party Punishment

In a first experiment, we manipulated the ascribed SES of the recipient of the unfair offers (i.e., Player B) while the SES of the allocator (i.e., Player A) was left unspecified. Results revealed an overall preference not to punish relative to non-punitive options, as indicated by a significant effect of the intercept,  $b=-22.975$ ,  $SE=0.750$ ,  $CI_{95\%}=[-24.445, -21.505]$ ,  $z=-30.633$ ,  $p<.001$ . We also observed significant main effects of recipient SES,  $b=3.196$ ,  $SE=0.588$ ,  $CI_{95\%}=[2.044, 4.347]$ ,  $z=5.439$ ,  $p<.001$ , and offer inequity,  $b=9.328$ ,  $SE=0.392$ ,  $CI_{95\%}=[8.560, 10.096]$ ,  $z=23.796$ ,  $p<.001$ . However, both effects were implicated in a significant  $SES \times$  Inequity interaction,  $b=-2.425$ ,  $SE=0.276$ ,  $CI_{95\%}=[-2.965, -1.884]$ ,  $z=-8.788$ ,  $p<.001$ . Analyses of simple effects (see Table 1) indicated that the preference for punishment increased as a function of offer inequity for both low- and high-SES recipients, but this effect of inequity was especially pronounced (i.e., 22.4% larger) for low-SES recipients. In fact, not only did highly

selfish offers made to low-SES recipients elicit the highest punishment rate, this also came the closest to eliminating the overall preference for non-punitive responses (see Fig 3A).

[Insert Fig 3]

**Fig 3. Percentage of decisions to punish by condition and experiment.** The recipient's and allocator's SES levels were manipulated in separate experiments (panels A and B, respectively). For ease of interpretation, percentages from the raw count data are plotted rather than fitted log odds ratios. For this reason, error bars are not included in this figure. In both experiments, the percentage of decisions to punish (vs. not punish) perpetrators of selfish allocations is plotted as a function of ascribed SES (i.e., high, low) and offer inequity (i.e., highly selfish, less selfish). In Experiment 1 (A:  $n=494$ ), we observed greater punishment of perpetrators who allocated selfish offers to low-SES victims. This was especially pronounced when offers were highly selfish. In Experiment 2 (B:  $n=505$ ), we observed greater punishment of high-SES (vs. low-SES) perpetrators, but only when offers were highly selfish. Significant simple effects of SES are indicated with solid lines; all simple effects of offer inequity (i.e., contrasts of same-color bars) were significant (see Tables 1 and 3). The dotted line represents equal preference for punitive and non-punitive options—all conditions were significantly different from this line (see Tables 2 and 4).

*Table 1. Simple Effects for Punitive versus Non-Punitive Decisions in Experiment 1.*

Contrast	Subset	<i>b</i>	<i>SE</i>	<i>CI</i> <sub>95%</sub>	<i>z</i>	<i>p</i>
High vs. Low Inequity	High SES	0.785	0.115	[0.560, 1.011]	6.829	< .001
	Low SES	0.878	0.113	[0.656, 1.099]	7.752	< .001
High vs. Low SES	High Inequity	-0.441	0.104	[-0.644, -0.237]	-4.245	< .001
	Low Inequity	-0.641	0.174	[-0.982, -0.300]	-3.686	< .001

*Table 2. Mean Preference for Punishment by Condition in Experiment 1.*

Offer Inequity	Recipient SES	<i>b</i>	<i>SE</i>	<i>CI</i> <sub>95%</sub>	<i>z</i>	<i>p</i>
High	High	-0.579	0.094	[-0.764, -0.394]	-6.146	< .001
	Low	-0.199	0.091	[-0.377, -0.022]	-2.203	.028
Low	High	-1.529	0.119	[-1.763, -1.295]	-12.830	< .001
	Low	-1.210	0.108	[-1.422, -0.998]	-11.200	< .001

*Note.* Negative betas reflect an overall preference for non-punitive solutions.

*Table 3. Simple Effects for Punitive versus Non-Punitive Decisions in Experiment 2.*

Contrast	Subset	<i>b</i>	<i>SE</i>	<i>CI</i> <sub>95%</sub>	<i>z</i>	<i>p</i>
High vs. Low Inequity	High SES	0.926	0.111	[0.709, 1.143]	8.378	< .001
	Low SES	0.936	0.124	[0.694, 1.179]	7.561	< .001
High vs. Low SES	High Inequity	0.603	0.094	[0.419, 0.787]	6.420	< .001
	Low Inequity	1.874	0.304	[1.279, 2.469]	6.173	< .001

*Table 4. Mean Preference for Punishment by Condition in Experiment 2.*

Offer Inequity	Allocator SES	<i>b</i>	<i>SE</i>	<i>CI</i> <sub>95%</sub>	<i>z</i>	<i>p</i>
High	High	0.245	0.090	[0.068, 0.421]	2.718	<b>.007</b>
	Low	-0.437	0.092	[-0.617, -0.258]	-4.772	<b>&lt; .001</b>
Low	High	-0.787	0.097	[-0.977, -0.598]	-8.147	<b>&lt; .001</b>
	Low	-1.513	0.116	[-1.740, -1.285]	-13.020	<b>&lt; .001</b>

*Note.* Negative betas reflect an overall preference for non-punitive solutions.

## Experiment 2: Effects of Allocator SES on Third-Party Punishment

In Experiment 1, we found that participants preferred to punish more if the allocator victimized a low-SES person than if the allocator victimized a high-SES person. Just as people seemed more concerned about Martin Shkreli's victimization of the disadvantaged (those depending on affordable prescription drugs) than those of privilege (the hedge fund investors who trusted him with their money), this finding suggests that participants were more averse to the financial exploitation of an already disadvantaged individual than they were to the exploitation of a relatively privileged individual. Independent of the recipient's SES, participants also preferred to punish more for increasingly selfish offers. These results, however, do not speak to the question of whether a similar SES-based bias exists for perpetrators of financial exploitation. Thus, in a second experiment, we manipulated the SES of the allocator (i.e., Player A). The SES of the recipient (i.e., Player B) was left unspecified.

Results revealed an overall preference not to punish versus punish, as indicated by a significant effect of the intercept,  $b = -26.276$ ,  $SE = 0.950$ ,  $CI_{95\%} = [-28.138, -24.414]$ ,  $z = -27.659$ ,  $p < .001$ . We also observed significant main effects of recipient SES,  $b = 7.979$ ,  $SE = 0.389$ ,  $CI_{95\%} = [7.217, 8.741]$ ,  $z = 20.516$ ,  $p < .001$ , and offer inequity,  $b = 8.187$ ,  $SE = 0.388$ ,  $CI_{95\%} = [7.426, 8.948]$ ,  $z = 21.094$ ,  $p < .001$ . However, both effects were implicated in a significant SES  $\times$  Inequity interaction,  $b = -1.214$ ,  $SE = 0.256$ ,  $CI_{95\%} = [-1.716, -0.711]$ ,  $z = -4.734$ ,  $p < .001$ . Analyses of simple effects (see Table 2) indicated that the preference for punishment increased as a function of offer

inequity for both low- and high-SES allocators, but this effect of inequity was especially pronounced (i.e., 17.5% larger) for high-SES allocators. In fact, not only did highly selfish offers made by high-SES allocators elicit the highest punishment rate, this was the only condition that significantly flipped the overall preference for non-punitive responses (see Fig 3B).

### **Comparison of Third-Party Punishment Across Experiments**

To determine whether manipulating the recipient's (vs. the allocator's) SES differentially affected decisions to punish selfish offers, we formally compared results across the experiments. To facilitate this comparison, we included experiment and all possible interactions with experiment as predictors in our analyses. We conducted an additional mixed-effects logistic regression predicting how decisions to punish selfish offers by experiment were predicted by both SES and offer inequity (see Methods for details).

As in the preceding analyses, the results revealed a prevailing preference for non-punitive options, as indicated by a significant effect of the intercept,  $b=-1.565$ ,  $SE=0.125$ ,  $CI_{95\%}=[-1.810, -1.319]$ ,  $z=-12.496$ ,  $p<.001$ . Results also indicated a greater overall tendency to punish in Experiment 2 than in Experiment 1,  $b=0.283$ ,  $SE=0.103$ ,  $CI_{95\%}=[0.082, 0.485]$ ,  $z=2.761$ ,  $p=.006$ . This increase in punitive decisions in Experiment 2 was driven by exchanges involving a high-SES individual (see Table 3 for simple effects statistics), as indicated by a significant interaction between target SES and experiment on punishment decisions,  $b=0.518$ ,  $SE=0.054$ ,  $CI_{95\%}=[0.413, 0.624]$ ,  $z=9.625$ ,  $p<.001$ . Tests of simple effects (Table 3) revealed that participants punished perpetrators with presumed higher standing relative to the victim more frequently. As depicted in Fig 4, we observed an increased preference to punish selfish allocators as a function of decreasing recipient SES (Experiment 1) and increasing allocator SES (Experiment 2).

[Insert Fig 4]

**Fig 4. Formal comparison of decisions to punish between experiments.** SES differentially shaped punishment decisions depending on whether SES varied for the recipient (Experiment 1) or the allocator (Experiment 2). For ease of interpretation, percentages from the raw count data are plotted rather than fitted log odds ratios. For this reason, error bars are not included in this figure. Independent of offer inequity, percentages of decisions to punish (vs. not punish) selfish offers are plotted as a function ascribed SES level (i.e., low, high) and experiment (i.e., Experiment 1, Experiment 2: total  $n=999$ ). Significant simple effects are indicated by distinct letters (i.e.,  $c > a > b$ ), all  $p < .05$ . See Table 5 for full contrast statistics. For all conditions plotted here, participants preferred non-punitive to punitive solutions, as indicated by all bars being significantly below the dotted line representing equal preference (see Table 6).

*Table 5. Simple Effects Between Experiments 2 versus 1 for Punitive versus Non-Punitive Options.*

Experiment 1	Experiment 2	<i>b</i>	<i>SE</i>	<i>CI</i> <sub>95%</sub>	<i>z</i>	<i>p</i>
High SES	High SES	0.700	0.108	[0.487, 0.912]	6.463	< .001
Low SES	Low SES	-0.244	0.103	[-0.445, -0.043]	-2.374	.018
Low SES	High SES	0.389	0.099	[0.196, 0.583]	3.943	< .001
High SES	Low SES	0.073	0.104	[-0.130, 0.277]	0.705	.481

*Table 6. Mean Preference for Punishment Across Experiments.*

Experiment	SES	<i>b</i>	<i>SE</i>	<i>CI</i> <sub>95%</sub>	<i>z</i>	<i>p</i>
Experiment 1	High	-1.915	0.247	[-2.399, -1.431]	-7.760	< .001
	Low	-1.279	0.182	[-1.636, -0.921]	-7.017	< .001
Experiment 2	High	-0.487	0.138	[-0.757, -0.216]	-3.526	< .001
	Low	-1.838	0.246	[-2.319, -1.357]	-7.487	< .001

*Note.* Negative betas reflect an overall preference for non-punitive solutions.

Additionally, we also observed a greater tendency to punish when financial offers involved a high-SES (vs. low-SES) party or when these offers were highly (vs. moderately) unfair, as indicated by significant main effects of SES,  $b=0.170$ ,  $SE=0.051$ ,  $CI_{95\%}=[0.071, 0.270]$ ,  $z=3.353$ ,  $p=.001$ , and offer inequity,  $b=1.016$ ,  $SE=0.072$ ,  $CI_{95\%}=[0.874, 1.158]$ ,  $z=14.051$ ,  $p<.001$ , respectively. All other effects were non-significant, ( $p>.52$ ).

## Discussion

The present findings show novel evidence in a controlled experimental context that social hierarchy can shape third-party perceptions of fairness even in a low-stakes financial allocation: Social hierarchy cues such as SES and the degree of inequity together influenced participants' decisions to punish privileged allocators who made selfish offers. These results reveal that perceived social hierarchy information has important consequences for how inequity is mitigated. In line with past research (21–23,65,85), participants tended to more frequently punish allocators who made especially selfish offers. However, when participants did opt to punish selfish allocators, this was especially likely when the victim was low in SES or if the allocator was high in SES. This is consistent with our initial predictions that participants would be more likely to punish (a) those who exploit the already disadvantaged, and (b) privileged perpetrators of inequity. Moreover, this tendency to favor the low-SES party in punishment decisions was especially pronounced in the context of highly selfish offers (but see Supplemental Text S2) and for low-SES participants (see Supplemental Text S6). Nonetheless, preferential treatment of low-SES (vs. high-SES) individuals was observed even for participants reporting relatively high SES (see Supplemental Text S6), suggesting that effects cannot be entirely due to ingroup favoritism. Another central finding from this study is that participants more frequently opted for punishment when perpetrator's SES was made explicit (viz., Experiment 2) rather than when the victim's SES was made explicit (viz., Experiment 1). Only when a high-SES perpetrator made a highly unfair offer did a majority of participants prefer punishment over the non-punitive options, suggesting that these privileged individuals elicited more desire for punishment when they exploited others to fully maximize their share.

505           Taken together, the present findings illustrate that status cues add an additional level of  
506 inequity to already inequitable exchanges that can bias subsequent decisions to punish, and that  
507 social hierarchy serves as a meaningful backdrop against which decisions about fairness are  
508 made. These findings add to the literature on punishment of fairness violations in hierarchical  
509 contexts (13,46–51). Previous work has so far focused largely on criminal offenders in an  
510 explicitly judicial context (e.g., 13,46,50,51), which may be biased by the salience of various  
511 stereotypes about criminality (52,53). In general, these studies illustrate that participants desire to  
512 punish the rich more severely for exploitative criminal violations. Consistent with these findings,  
513 the present study also finds that high-SES perpetrators are punished more severely for their  
514 selfish offers. However, unlike these previous studies, this was found to be the case even in a  
515 context that minimized the salience of punishment—namely, in the absence of an explicitly  
516 judicial context or illegal behavior and when non-punitive means of restoring equity were  
517 available. In sum, the present findings rule out the possibility that preferences for punishing  
518 selfishness among the privileged are driven by ascribed illegal behavior in an explicitly judicial  
519 context where punitive outcomes are salient. Instead, punishing the privileged for selfishness  
520 may reflect social norms (e.g., *noblesse oblige*) that transcend contexts and cultures (15).

521           Other previous work suggests that encouraging people to think of the rich as having it all  
522 and/or the poor as miserable makes them more amenable to thinking about justice (42,67). The  
523 present study goes beyond merely thinking about justice, ultimately providing evidence that  
524 salient examples of widening inequity between two individuals who are already unequal by  
525 virtue of their SES trigger punitive corrective action from third-party arbitrators. In Experiment  
526 1, we observed an inclination to punish allocators who withheld money from an already  
527 economically disadvantaged low-SES victim. Moreover, participants were especially willing to

punish a selfish allocator if that individual was ascribed a privileged, high-SES background (Experiment 2), which is consistent with work documenting our sensitivity to social rank (4–6,39) and egalitarian tendencies (1,2). In other words, participants' decisions to punish appear to reflect an aversion to concrete instances of widening inequality, both when the poor do not receive a fair share, but especially when the rich get richer at the expense of others. Although plausible, it remains to be seen whether this aversion is exacerbated when the status of both the perpetrator and victim is made salient.

These findings have important implications for judicial decision making in cases involving inequity in financial exchanges. Making the relative SES of plaintiffs and defendants explicit may bias jurors in favor of the party of lower SES. Although this appears to run counter to the principle that justice is blind to characteristics of parties involved in a given case, knowledge of the SES of plaintiffs/defendants may be particularly germane to decisions about the relative harm of financial inequity. Barring that information from the court may hinder the court's ability to accurately assess the psychological and material harms of financial exploitation. For example, a monetary loss is likely to elicit greater distress and financial hardship in a low-SES individual than in a high-SES individual.

The present work helps to differentiate between two possible accounts of how SES may shape third-party fairness decisions. A greater willingness to allow exploitation of the rich to go unpunished (see Experiment 1) is inconsistent with predictions stemming from lay beliefs that the rich are relatively law-abiding (86) and less accustomed to suffering (87) in comparison to the poor. Such beliefs would presumably make the victimization of high-SES individuals potentially more salient and deserving of punishment (61,64). However, we instead observed greater punishment on behalf of low-SES compared to high-SES victims in Experiment 1.



551 Additionally, beliefs that the rich are relatively good and less accustomed to suffering have been  
552 linked to more lenient punishments for rich perpetrators (88,89), which is the opposite of what  
553 we observed in Experiment 2. Instead, participants' preferences to punish selfish high-SES  
554 allocators and protect low-SES recipients by punishing those who exploit them are more  
555 consistent with (a) violation of the norm that the rich should be generous with or at least not  
556 exploit the poor (15), and (b) work showing that the rich effectively pursue their own self-  
557 interest whereas the poor are generally perceived as well-intentioned, albeit less effective  
558 (46,67,90). These latter group-based beliefs are tied to distinct motivated emotions, with people  
559 experiencing envy of the privileged and pity for those who are deemed well-intentioned but  
560 disadvantaged (90,91). Although the present study did not directly examine participant emotions,  
561 previous work has implicated a significant role of affect in justice-related decision making.  
562 Particularly for envied classes, individuals show greater evidence of pleasure when members of  
563 these classes experience misfortune (68). It is therefore possible that participants enjoyed  
564 punishing high SES allocators. A similar mechanism may also underlie the tendency for  
565 participants to favor punishment as a means of correcting the exploitation of low-SES recipients.  
566 Related to this possibility, one fMRI study showed that punishment (vs. compensatory  
567 redistribution) was associated with greater activity in brain regions involved in reward  
568 processing (23). Accordingly, future work may explore whether the SES of the  
569 allocator/recipient may accentuate third-party preferences for punishment vis-a-vis associated  
570 emotional and reward responses.

571         It will be important for future work to explore whether the present findings may hold  
572 when more is at stake (e.g., greater sums of money) and/or after having met the individuals  
573 involved in the Justice Game exchanges in real life. For example, with a few exceptions (63),

sociological data often show that wealthy (vs. poor) victims elicit greater, not lesser, penalties for criminal offenders (61,64). One exciting future direction at the rich but often-overlooked nexus of sociology and social psychology (92) would be to explore when and how structural factors (e.g., aspects of the legal/judicial system that privilege the wealthy) can be overcome by the individual-level aversion to widening inequality evinced by the present findings. In other words, how can the presently reported effects of SES on perceived justice in individual interactions scale up to the society level? Americans (among other nationalities) tend to underestimate the magnitude of inequality at the level of society (4,45,93). This misperception combined with ideological mobility beliefs (5,12,94), perceived competition between racial/ethnic groups (41,94–96) and dissatisfaction with the government (97,98), can diminish concern about the negative impacts of growing inequality (99), resulting in weak support for redistributive policies (9,12,94,97,98). One means of catalyzing efforts to reduce inequality may be to make the issue more concrete. In the aftermath of the 2009 stock market crash, the Occupy Wall Street movement emerged under the slogan, “We are the 99%”. On the one hand, this slogan encapsulated in a few words ongoing inequities in American society where the wealthy elite wield a disproportionate share of capital and political influence relative to the poor and middle class (100). On the other hand, this slogan was relatively abstract, making it easy for critics to misconstrue (101). One line of research suggests that focusing on individuals (e.g., CEOs) rather than groups of people (e.g., corporations) may make punishments of fairness violations more appealing (102). Accordingly, one means of increasing support for redistributive policies (punitive or non-punitive) may be to frame inequality as an interaction between two people: one representing the “haves” (i.e., the allocators) and the other representing the “have nots” (i.e., the recipients) within the broader American social hierarchy. In sum, the present findings highlight

## Social status in justice decisions

- 597 the important role of social hierarchy in the perception of fairness, with critical implications for
- 598 real-world retributive and restorative justice.

599

## References

- 600 1. Henrich J, Boyd R, Bowles S, Camerer C, Fehr E, Gintis H, et al. In search of homo  
601 economicus: Behavioral experiments in 15 small-scale societies. *Am Econ Rev.*  
602 2001;91:73–8.
- 603 2. Herrmann B, Thöni C, Gächter S. Altruistic punishment across societies. *Science* (80- ).  
604 2008;319:1362–7.
- 605 3. Fehr E, Fischbacher U. Third-party punishment and social norms. *Evol Hum Behav.*  
606 2004;25:63–87.
- 607 4. Norton MI, Ariely D. Building a better America—One wealth quintile at a time. *Perspect*  
608 *Psychol Sci.* 2011;6:9–12.
- 609 5. Davidai S, Gilovich T. Building a more mobile America — One income quintile at a time.  
610 *Perspect Psychol Sci.* 2015;10:60–71.
- 611 6. Kiatpongsan S, Norton MI. How much (more) should CEOs make? A universal desire for  
612 more equal pay. *Perspect Psychol Sci.* 2014;9:587–93.
- 613 7. Jost JT, Hunyady O. Antecedents and consequences of system-justifying ideologies. *Curr*  
614 *Dir Psychol Sci.* 2005;14:114–8.
- 615 8. Hauser OP, Hilbe C, Chatterjee K, Nowak MA. Social dilemmas among unequals. *Nature.*  
616 2019;572(7770):524–7.
- 617 9. Sands ML. Exposure to inequality affects support for redistribution. *Proc Natl Acad Sci.*  
618 2017;114:663–8.
- 619 10. Hackel LM, Zaki J. Propagation of economic inequality through reciprocity and  
620 reputation. *Psychol Sci.* 2018;29:604–13.
- 621 11. Kraus MW, Keltner D. Social class rank, essentialism, and punitive judgment. *J Pers Soc*

Psychol. 2013;105:247–61.

12. Wakslak CJ, Jost JT, Tyler TR, Chen ES. Moral outrage mediates the dampening effect of system justification on support for redistributive social policies. *Psychol Sci.* 2007;18:267–74.
13. Cullen FT, Hartman JL, Jonson CL. Bad guys: Why the public supports punishing white-collar offenders. *Crime, Law Soc Chang.* 2009;51:31–44.
14. Hauser O, Kraft-Todd G, Rand D, Nowak M. Invisible inequality leads to punishing the poor and rewarding the rich. *Behav Public Policy.* 2019;
15. Fiddick L, Cummins DD, Janicki M, Lee S, Erlich N. A cross-cultural study of noblesse oblige in economic decision-making. *Hum Nat.* 2013;24:318–35.
16. Long C, Hays T. “Pharma Bro” Martin Shkreli cries in court, is sentenced to 7 years for securities fraud. *Chicago Tribune [Internet].* 2018 Mar 9; Available from: <https://www.chicagotribune.com/business/ct-biz-martin-shkreli-sentencing-20180309-story.html>
17. Thomas Z, Swift T. Who is Martin Shkreli - “the most hated man in America”? *BBC News [Internet].* 2017 Aug 4; Available from: <https://www.bbc.com/news/world-us-canada-34331761>
18. Public enemy [Internet]. *Harper’s Magazine.* 2017. Available from: <https://harpers.org/archive/2017/09/public-enemy/>
19. McCall L, Burk D, Laperrière M, Richeson JA. Exposure to rising inequality shapes Americans’ opportunity beliefs and policy support. *Proc Natl Acad Sci.* 2017;114:9593–8.
20. Ingraham C. How rising inequality hurts everyone, even the rich. *The Washington Post [Internet].* 2018; Available from:

- 645 [https://www.washingtonpost.com/news/wonk/wp/2018/02/06/how-rising-inequality-hurts-](https://www.washingtonpost.com/news/wonk/wp/2018/02/06/how-rising-inequality-hurts-everyone-even-the-rich/?noredirect=on)  
646 [everyone-even-the-rich/?noredirect=on](https://www.washingtonpost.com/news/wonk/wp/2018/02/06/how-rising-inequality-hurts-everyone-even-the-rich/?noredirect=on)
- 647 21. FeldmanHall O, Sokol-Hessner P, Van Bavel JJ, Phelps EA. Fairness violations elicit  
648 greater punishment on behalf of another than for oneself. *Nat Commun.* 2014;5(5306).
- 649 22. FeldmanHall O, Otto AR, Phelps EA. Learning moral values: Another's desire to punish  
650 enhances one's own punitive behavior. *J Exp Psychol Gen.* 2018;147:1211–24.
- 651 23. Stallen M, Rossi F, Heijne A, Smidts A, De Dreu CKW, Sanfey AG. Neurobiological  
652 mechanisms of responding to injustice. *J Neurosci* [Internet]. 2018;38:1242–1217.  
653 Available from: <http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.1242-17.2018>
- 654 24. Rabellino D, Morese R, Ciaramidaro A, Bara BG, Bosco FM. Third-party punishment:  
655 Altruistic and anti-social behaviours in in-group and out-group settings. *J Cogn Psychol.*  
656 2016;28:486–95.
- 657 25. Lo Gerfo E, Gallucci A, Morese R, Vergallito A, Ottone S, Ponzano F, et al. The role of  
658 ventromedial prefrontal cortex and temporo-parietal junction in third-party punishment  
659 behavior. *Neuroimage.* 2019;200:501–10.
- 660 26. Henrich J, Ensminger J, McElreath R, Barr A, Barrett C, Bolyanatz A, et al. Markets,  
661 religion, community size, and the evolution of fairness and punishment. *Science* (80- )  
662 [Internet]. 2010;1480:1480–4. Available from:  
663 <http://www.ncbi.nlm.nih.gov/pubmed/20299588>
- 664 27. Boehm C. Ancestral hierarchy and conflict. *Science* (80- ) [Internet]. 2012;336:844–7.  
665 Available from: <http://www.sciencemag.org/content/336/6083/844.short>
- 666 28. Brosnan SF, de Waal FBM. Fairness in animals: Where to from here? *Soc Justice Res.*  
667 2012;25:336–51.

- 668 29. Kraus MW, Piff PK, Keltner D. Social class as culture: The convergence of resources and  
669 rank in the social realm. *Curr Dir Psychol Sci.* 2011;20:246–50.
- 670 30. Magee JC, Galinsky AD. Social hierarchy: The self reinforcing nature of power and  
671 status. *Acad Manag Ann.* 2008;2:351–98.
- 672 31. Rucker DD, Galinsky AD, Magee JC. The Agentic–Communal Model of Advantage and  
673 Disadvantage: How inequality produces similarities in the psychology of power, social  
674 class, gender, and race. In: Olson JM, editor. *Advances in Experimental Social*  
675 *Psychology.* Cambridge, MA: Academic Press; 2018. p. 71–125.
- 676 32. Steckler CM, Tracy JL. The emotional underpinnings of social status. In: Cheng JT, Tracy  
677 JL, Anderson C, editors. *The psychology of social status.* New York, NY: Springer; 2014.  
678 p. 201–24.
- 679 33. Tracy JL, Shariff AF, Zhao W, Henrich J. Cross-cultural evidence that the nonverbal  
680 expression of pride is an automatic status signal. *J Exp Psychol Gen.* 2013;142:163–80.
- 681 34. Kraus MW, Park JW, Tan JJX. Signs of social class: The experience of economic  
682 inequality in everyday life. *Perspect Psychol Sci.* 2017;12:422–35.
- 683 35. Fiske ST, Dupree CH, Nicolas G, Swencionis JK. Status, power, and intergroup relations:  
684 The personal is the societal. *Curr Opin Psychol* [Internet]. 2016;11:44–8. Available from:  
685 <http://dx.doi.org/10.1016/j.copsyc.2016.05.012>
- 686 36. Cheng JT, Tracy JL, Foulsham T, Kingstone A, Henrich J. Two ways to the top: Evidence  
687 that dominance and prestige are distinct yet viable avenues to social rank and influence. *J*  
688 *Pers Soc Psychol.* 2013;104:103–25.
- 689 37. Mattan BD, Wei KY, Cloutier J, Kubota JT. The social neuroscience of race-based and  
690 status-based prejudice. *Curr Opin Psychol.* 2018;24:27–34.

- 691 38. Fiske ST. Interpersonal stratification: Status, power, and subordination. In: Fiske ST,  
692 Gilbert DT, Lindzey G, editors. *Handbook of social psychology*. Hoboken, NJ: Wiley;  
693 2010. p. 941–82.
- 694 39. Mattan BD, Kubota JT, Cloutier J. How social status shapes person perception and  
695 evaluation: A social neuroscience perspective. *Perspect Psychol Sci*. 2017;12:468–507.
- 696 40. Brown-Iannuzzi JL, Lundberg KB, Kay AC, Payne BK. Subjective status shapes political  
697 preferences. *Psychol Sci*. 2015;26:15–26.
- 698 41. Wilkins CL, Wellman JD, Babbitt LG, Toosi NR, Schad KD. You can win but I can't  
699 lose: Bias against high-status groups increases their zero-sum beliefs about discrimination.  
700 *J Exp Soc Psychol* [Internet]. 2015;57:1–14. Available from:  
701 <http://dx.doi.org/10.1016/j.jesp.2014.10.008>
- 702 42. Bruckmüller S, Reese G, Martiny SE. Is higher inequality less legitimate? Depends on  
703 how you frame it! *Br J Soc Psychol*. 2017;56:766–81.
- 704 43. Brown-Iannuzzi JL, Cooley E, McKee SE, Hyden C. Wealthy Whites and poor Blacks:  
705 Implicit associations between racial groups and wealth predict explicit opposition toward  
706 helping the poor. *J Exp Soc Psychol* [Internet]. 2019;82:26–34. Available from:  
707 <https://doi.org/10.1016/j.jesp.2018.11.006>
- 708 44. Moore-Berg SL, Karpinski A. An intersectional approach to understanding how race and  
709 social class affect intergroup processes. *Soc Personal Psychol Compass*.  
710 2019;13(1):e12426.
- 711 45. Kraus MW, Onyeador IN, Daumeyer NM, Rucker JM, Richeson JA. The misperception of  
712 racial economic inequality. *Perspect Psychol Sci*. 2019;
- 713 46. Fragale AR, Rosen B, Xu C, Merideth I. The higher they are, the harder they fall: The



- effects of wrongdoer status on observer punishment recommendations and intentionality attributions. *Organ Behav Hum Decis Process* [Internet]. 2009;108:53–65. Available from: <http://dx.doi.org/10.1016/j.obhdp.2008.05.002>
47. Ball S, Eckel CC. The economic value of status. *J Socio Econ*. 1998;27:495–514.
48. Hong K, Bohnet I. Status and distrust: The relevance of inequality and betrayal aversion. *J Econ Psychol*. 2007;28:197–213.
49. Eisenegger C, Naef M, Snozzi R, Heinrichs M, Fehr E. Prejudice and truth about the effect of testosterone on human bargaining behaviour. *Nature* [Internet]. 2010;463:356–9. Available from: <http://dx.doi.org/10.1038/nature08711>
50. Redford L, Ratliff KA. Retribution as hierarchy regulation: Hierarchy preferences moderate the effect of offender socioeconomic status on support for retribution. *Br J Soc Psychol*. 2018;57:75–94.
51. Redford L, Ratliff KA. Pride and punishment: Entitled people’s self-promoting values motivate hierarchy-restoring retribution. *Eur J Soc Psychol*. 2018;48:303–19.
52. Bodenhausen G V., Wyer RS. Effects of stereotypes on decision making and information-processing strategies. *J Pers Soc Psychol*. 1985;48:267–82.
53. Macrae CN, Shepherd JW. Do criminal stereotypes mediate juridic judgements? *Br J Soc Psychol*. 1989;28:189–91.
54. Lei RF, Bodenhausen G V. Racial assumptions color the mental representation of social class. *Front Psychol* [Internet]. 2017;8(519). Available from: <http://journal.frontiersin.org/article/10.3389/fpsyg.2017.00519/full>
55. Penner AM, Saperstein A. How social status shapes race. *Proc Natl Acad Sci*. 2008;105:19628–30.

56. Freeman JB, Penner AM, Saperstein A, Scheutz M, Ambady N. Looking the part: Social status cues shape race perception. *PLoS One* [Internet]. 2011;6(9):e25107. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3180382&tool=pmcentrez&rendertype=abstract>
57. Gino F, Pierce L. Robin Hood under the hood: Wealth-based discrimination in illicit customer help. *Organ Sci*. 2010;21:1176–94.
58. Steffensmeier D, Ulmer J, Kramer J. The interaction of race, gender, and age in criminal sentencing: The punishment cost of being young, black, and male. *Criminology*. 1998;36:763–98.
59. Mitchell TL, Haw RM, Pfeifer JE, Meissner C a. Racial bias in mock juror decision-making: a meta-analytic review of defendant treatment. *Law Hum Behav* [Internet]. 2005 Dec [cited 2014 Aug 16];29(6):621–37. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16382353>
60. Myers MA, Talarico SM. Urban justice, rural injustice? Urbanization and its effect on sentencing. *Criminology*. 1986;24:367–91.
61. Phillips S. Status disparities in the capital of capital punishment. *Law Soc Rev*. 2009;43:807–38.
62. Buckler KG. Supreme Court outcomes in criminal justice cases (1994-2012 terms): An examination of status differential and amici curiae effects. *Crim Justice Policy Rev*. 2015;26:773–804.
63. Benson ML, Walker E. Sentencing the white-collar offender. *Am Sociol Rev*. 1988;53:294–302.

64. Wheeler S, Weisburd D, Bode N. Sentencing the white-collar offender: Rhetoric and reality. *Am Sociol Rev.* 1982;47:641–59.
65. Son J-Y, Bhandari A, FeldmanHall O. Crowdsourcing punishment: Individuals reference group preferences to inform their own punitive decisions. *Sci Rep.* 2019;9(11625).
66. Fiske ST, Cuddy AJC, Glick P, Xu J. A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *J Pers Soc Psychol.* 2002;82:878–902.
67. Kay AC, Jost JT. Complementary justice: Effects of “poor but happy” and “poor but honest” stereotype exemplars on system justification and implicit activation of the justice motive. *J Pers Soc Psychol* [Internet]. 2003;85:823–37. Available from: <http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-3514.85.5.823>
68. Cikara M, Fiske ST. Stereotypes and schadenfreude: Affective and physiological markers of pleasure at outgroup misfortunes. *Soc Psychol Personal Sci.* 2012;3:63–71.
69. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2019. Available from: <https://www.r-project.org/>
70. Bates D, Maechler B, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw.* 2015;67:1–48.
71. Horton JJ, Rand DG, Zeckhauser RJ. The online laboratory: Conducting experiments in a real labor market. *Exp Econ.* 2011;14:399–425.
72. Mason W, Suri S. Conducting behavioral research on Amazon’s Mechanical Turk. *Behav Res Methods* [Internet]. 2012;44:1–23. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21717266%5Cnhttp://www.springerlink.com/index/>

- 10.3758/s13428-011-0124-6
73. Paolacci G, Chandler J, Ipeirotis PG. Running experiments on Amazon Mechanical Turk. *Judgm Decis Mak* [Internet]. 2010;5:411–9. Available from: <http://repub.eur.nl/pub/31983>
74. Buhrmester M, Kwang T, Gosling SD. Amazon’s Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspect Psychol Sci* [Internet]. 2011;6:3–5. Available from: <http://www.sas.upenn.edu/~baron/journal/8210/jdm8210.pdf>
75. Ma DS, Correll J, Wittenbrink B. The Chicago face database: A free stimulus set of faces and norming data. *Behav Res Methods* [Internet]. 2015;47:1122–35. Available from: <http://link.springer.com/10.3758/s13428-014-0532-5>
76. Oh D, Shafir E, Todorov A. Economic status cues from clothes affect perceived competence from faces. *Nat Hum Behav*. 2019;
77. Fiske ST, Cuddy AJC, Glick P. Universal dimensions of social cognition: Warmth and competence. *Trends Cogn Sci*. 2007;11:77–83.
78. Mattan BD, Kubota JT, Li T, Venezia SA, Cloutier J. Implicit evaluative biases toward targets varying in race and socioeconomic status. *Personal Soc Psychol Bull*. 2019;45:1512–27.
79. Mattan BD, Kubota JT, Dang TP, Cloutier J. External motivation to avoid prejudice alters neural responses to targets varying in race and status. *Soc Cogn Affect Neurosci*. 2018;13:22–31.
80. Mattan BD, Kubota JT, Li T, Dang TP, Cloutier J. Motivation modulates brain networks in response to faces varying in race and status: A multivariate approach. *eNeuro*. 2018;5.
81. Gyurovski I, Kubota JT, Cardenas-Iniguez C, Cloutier J. Social status level and dimension interactively influence person evaluations indexed by P300s. *Soc Neurosci* [Internet].

- 2017;13:333–45. Available from: <https://doi.org/10.1080/17470919.2017.1326400>
82. Cloutier J, Gyurovski I. Ventral medial prefrontal cortex and person evaluation: Forming impressions of others varying in financial and moral status. *Neuroimage*. 2014 Jun 14;100:535–43.
83. Bates D, Kliegl R, Vasishth S, Baayan H. Parsimonious mixed models [Internet]. 2015. Available from: <https://arxiv.org/pdf/1506.04967v1.pdf>
84. Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang* [Internet]. 2013;68:255–78. Available from: <http://dx.doi.org/10.1016/j.jml.2012.11.001>
85. Morese R, Rabellino D, Sambataro F, Perussia F, Valentini MC, Bara BG, et al. Group membership modulates the neural circuitry underlying third party punishment. *PLoS One*. 2016;11(11):e0166357.
86. Varnum MEW. What are lay theories of social class? *PLoS One*. 2013;8(7):e70589.
87. Hoffman KM, Trawalter S. Assumptions about life hardship and pain perception. *Gr Process Intergr Relations*. 2016;19:493–508.
88. Reiman J, Leighton P. *The Rich Get Richer and the Poor Get Prison*. 10th ed. New York, NY: Routledge; 2016.
89. Freeman NJ. Socioeconomic status and belief in a just world: Sentencing of criminal defendants. *J Appl Soc Psychol*. 2006;36:2379–94.
90. Cuddy AJC, Fiske ST, Glick P. Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. In: Zanna MP, editor. *Advances in Experimental Social Psychology* [Internet]. Amsterdam, The Netherlands: Academic Press; 2008. p. 61–149. Available from:

<http://linkinghub.elsevier.com/retrieve/pii/S0065260107000020>

91. Sandstrom GM, Schmader T, Croft A, Kwok N. A social identity threat perspective on being the target of generosity from a higher status other. *J Exp Soc Psychol* [Internet]. 2019;82:98–114. Available from: <https://doi.org/10.1016/j.jesp.2018.12.004>
92. Oishi S, Kesebir S, Snyder BH. Sociology: A lost connection in social psychology. *Personal Soc Psychol Rev*. 2009;13:334–53.
93. Kraus MW, Rucker JM, Richeson JA. Americans misperceive racial economic equality. *Proc Natl Acad Sci*. 2017;114:10324–31.
94. Wilkins CL, Kaiser CR. Racial progress as threat to the status hierarchy: Implications for perceptions of anti-white bias. *Psychol Sci*. 2014;25:439–46.
95. Krosch AR, Amodio DM. Economic scarcity alters the perception of race. *Proc Natl Acad Sci* [Internet]. 2014;111:9079–84. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24927595>
96. Krosch AR, Tyler TR, Amodio DM. Race and recession: The effect of economic scarcity and egalitarian motivation on racial discrimination. *J Pers Soc Psychol* [Internet]. 2017;113:892–909. Available from: <http://doi.apa.org/getdoi.cfm?doi=10.1037/pspi0000112>
97. Macdonald D. Trust in government and the American public's responsiveness to rising inequality. *Polit Res Q*. 2019;
98. Wright G. The political implications of American concerns about economic inequality. *Polit Behav*. 2018;40:321–43.
99. Wilkinson RG, Pickett KE. Income inequality and social dysfunction. *Annu Rev Sociol* [Internet]. 2009;35:493–511. Available from:

- 852            <http://www.annualreviews.org/doi/abs/10.1146/annurev-soc-070308-115926>
- 853    100.    Krugman P. We are the 99.9% [Internet]. The New York Times. 2011 [cited 2018 Dec
- 854            19]. Available from: <https://www.nytimes.com/2011/11/25/opinion/we-are-the-99-9.html>
- 855    101.    Memmott M. For those who aren't fans of the "99 percent," there's the "53 percent"
- 856            [Internet]. The Two-Way: Breaking News from NPR. 2011 [cited 2018 Dec 19].
- 857            Available from: [https://www.npr.org/sections/thetwo-way/2011/10/11/141239358/for-](https://www.npr.org/sections/thetwo-way/2011/10/11/141239358/for-those-who-arent-fans-of-the-99-percent-theres-the-53-percent)
- 858            [those-who-arent-fans-of-the-99-percent-theres-the-53-percent](https://www.npr.org/sections/thetwo-way/2011/10/11/141239358/for-those-who-arent-fans-of-the-99-percent-theres-the-53-percent)
- 859    102.    Tang S, Gray K. CEOs imbue organizations with feelings, increasing punishment
- 860            satisfaction and apology effectiveness. *J Exp Soc Psychol.* 2018;79:115–25.
- 861

862 **Supporting Information Captions**

863 **Text S1. Justice Game design and instructions.**

864 **Text S2. Preferences for redistribution over acceptance of selfish offers.**

865 **Text S3. All analyses of hyper-generous offers.**

866 **Text S4. Exploratory measures and demographic items.**

867 **Text S5. Standard model optimization procedure.**

868 **Text S6. Exploratory analyses of subjective SES.**