Moral Judgment Is Sensitive to Bargaining Power

Arthur Le Pargneux[1, 2] and Fiery Cushman[2]
[1] Behavioural Science Group, Warwick Business School, University of Warwick
[2] Department of Psychology, Harvard University

For contractualist accounts of morality, actions are moral if they correspond to what rational or reasonable agents would agree to do, were they to negotiate explicitly. This, in turn, often depends on each party's bargaining power, which varies with each party's stakes in the potential agreement and available alternatives in case of disagreement. If there is an asymmetry, with one party enjoying higher bargaining power than another, this party can usually get a better deal, as often happens in real negotiations. A strong test of contractualist accounts of morality, then, is whether moral judgments do take bargaining power into account. We explore this in five preregistered experiments ($n = 3,025$; U.S.-based Prolific participants). We construct scenarios depicting everyday social interactions between two parties in which one of them can perform a mutually beneficial but unpleasant action. We find that the same actions (asking the other to perform the unpleasant action or explicitly refusing to do it) are perceived as less morally appropriate when performed by the party with lower bargaining power, as compared to the party with higher bargaining power. In other words, participants tend to give more moral leeway to parties with better bargaining positions and to hold disadvantaged parties to stricter moral standards. This effect appears to depend only on the relative bargaining power of each party but not on the magnitude of the bargaining power asymmetry between them. We discuss implications for contractualist theories of moral cognition and the emergence and persistence of unfair norms and inequality.

Public Significance Statement
Many social interactions involve opportunities for mutual benefit. By engaging in negotiation— sometimes explicitly, but often tacitly—we decide what each party should do and enter arrangements that we anticipate will be advantageous for everyone involved. Contractualist theories of morality insist on the fundamental role played by such bargaining procedures in determining what constitutes appropriate and inappropriate behavior. But the outcome of a negotiation often depends on each party's bargaining power and their relative positions if an agreement cannot be reached. And situations in which each party enjoys equal bargaining power are rare. Here, we investigate the influence of bargaining power on our moral judgments. Consistent with contractualist accounts, we find that moral judgments take bargaining power considerations into account, to the benefit of the powerful party, and that parties with lower bargaining power are held to stricter moral standards.

Keywords: moral judgment, bargaining power, contractualism, inequality

Supplemental materials: https://doi.org/10.1037/xge0001678.supp

Recently, there has been growing interest in "contractualist" models of moral cognition. These assert that moral judgments track the agreements that people would come to—ones that often achieve mutual benefit. Sometimes this is because cultural or biological adaptive processes converge on bargaining solutions (André et al., 2023; Baumard, 2016; Baumard et al., 2013; Binmore, 1994a, 1994b, 2005). Other times, this is because people actually negotiate the norms that govern their community (Ostrom, 1990, 2000; Ostrom et al., 1992). And, other times, it is because people intelligently model the kinds of agreements that would be likely if there were a negotiation (Levine, Chater, et al., 2024)—a process sometimes called "virtual bargaining" (Chater et al., 2022; Misyak et al., 2014). In each case, contractualist accounts of moral cognition argue that its central function is to approximate the results of explicit negotiation among relevant stakeholders seeking mutual advantage (André et al., 2023; Levine, Chater, et al., 2024).

Prior empirical work testing the predictions of the contractualist approach (e.g., Everett et al., 2016; Le Pargneux et al., 2024; Levine, Kleiman-Weiner, et al., 2024; Levine et al., 2020) mostly focus on the agreements that people come to in symmetrical relationships, where neither party holds any particular advantage. Yet, the outcome of a negotiation often depends on each party's bargaining power (BP), which can often be determined by looking at what each of them stands to gain from the bargain, and what their best alternative is if they fail to reach an agreement. A classic example is the "buyer's market" versus the "seller's market" and its effect on the perceived fair value of a home. Under a contractualist lens, bargaining power considerations are expected to influence moral judgments.[1] Here, we ask whether this is the case. Put simply: Do people think that it is morally appropriate for those with stronger negotiating positions to get more beneficial arrangements?

Many of our social interactions are indeed asymmetrical, with one party enjoying higher bargaining power than another. In relationships between employers and job applicants, managers and subordinates, parents and children, athletes and sports teams, or two romantic partners, bargaining power can shape what each party will be willing to agree to, their role in the interaction, and how they will behave (Bacharach & Lawler, 1981; Emerson, 1962). What constitutes bargaining power? A useful place to start is to think of it in terms of "the ability to secure another's agreement on one's own terms" (Kuhn et al., 1983, p. 143). More formally, according to power-dependence theory, a central account in the negotiations literature, "the power of A over B is equal to and based upon the dependence of B upon A" (Emerson, 1962, p. 33). In the context of bargaining, *dependence* can be divided into two main components: *outcome value* and *available alternatives*. First, the dependence of B upon A is "proportional to the value attributed to the outcome at stake" by B. Second, it is "inversely proportional to the availability of the same or better outcomes from alternative sources" to B (Kim & Fragale, 2005, p. 73), that is, B's "best alternative to a negotiated agreement" or BATNA (Pinkley et al., 1994). Thus, the bargaining power of each party depends both on what is at stake for each of them (the "importance of" and "need" for the outcome of each bargainer, Bacharach & Lawler, 1981, p. 220) and on the quality of their alternatives. Importantly, *outcome value* (henceforth "stakes") and *available alternatives* have a mutually dependent role in determining bargaining power. The closer the value of a bargain to the value of its alternatives, the stronger the bargaining power of a party.

Consider a variant of an example discussed by Bruner and O'Connor (2016). A graduate student and her advisor are coauthoring

a scientific article. Quick publication in a peer-reviewed journal is crucial for the student, as it can drastically improve her employment prospects. For her advisor, a tenured professor publishing several articles each year, the timing of publication makes little difference. One of them has to handle the submission, which involves a number of tedious and time-consuming tasks including proofreading, formatting, double checking references, and completing multiple forms on the publisher's submission portal. Who should do it? Intuitively, it seems that the graduate student, who has more to lose if the article is not published, should take the initiative. Asking her advisor to handle the submission seems misplaced and explicitly refusing to do so would be downright inappropriate. But for the advisor with higher bargaining power, making the same request seems perfectly acceptable and explicitly refusing to handle the submission is an option.[2] This suggests that our intuitions about appropriate and inappropriate conduct—and thus our moral judgments—might be influenced by inferences about the relative bargaining position of each party involved in an interaction.

Yet, this prediction is at odds with various widely held conceptions of justice that emphasize proportional treatment according to merit (Aristotle, c. 350 BC/2009), the importance of egalitarian concerns and treating everyone equally (Nagel, 1995) or of not worsening the situation of the least advantaged (Rawls, 1971). If both authors have made equal contributions to the article, it might be considered unfair for the disadvantaged party (the student) to be held to stricter moral standards than the advantaged party (the advisor). By analogy, we do not expect the poor to pay more taxes than the rich because they are less powerful. It seems particularly puzzling that the graduate student, who is already in a worse position overall, should be further disadvantaged, simply because of her bargaining position.

Although empirical investigations of bargaining power are mostly absent in moral psychology, some existing work in experimental economics lends support to the contractualist view. These studies investigate the role of power (broadly construed) on behavior in modifications of standard economic games (ultimatum, trust, public goods games, etc.), as well as its effect on fairness perceptions. Although there is variability in the results and in the way in which power is operationalized, the main findings can be summarized as follows. First, monetary divisions tend to reflect power imbalances between the players: Players endowed with more power contribute less to common pools (Hsu, 2008), increase their demands (Hennig-Schmidt et al., 2018), and keep more money for themselves (Amasino et al., 2023; Lois & Riedl, 2022; Rode & Le Menestrel, 2011; Rustichini & Villeval, 2014). Second, fairness perceptions of monetary divisions also seem to track power asymmetries. For example, advantaged and disadvantaged players perceive as fair

---

[1] Importantly, understood in its broad sense, contractualism (like consequentialism, deontology, or virtue ethics) comes in a variety of flavors. For some theories (in particular psychological accounts inspired by the Hobbesian *contractarian* strand), bargaining power is expected to play an important *descriptive* role. This is not the case for other influential accounts (e.g., *normative* theories from the Kantian *contractualist* tradition) which posit equal bargaining positions for the contractors based on certain commitments to impartiality and equality.

[2] To keep things simple, we assume that additional considerations (e.g., first authorship, time spent on the article, resources invested) cannot easily clarify who should handle the submission in this case, that is, first authorship is shared, both parties have worked equally hard on the article, and so forth. In addition, the asymmetry in seniority is not crucial here, equivalent examples in which both authors are of equal rank (e.g., graduate students, postdocs, professors) can easily be imagined.

various deviations from equal sharing that favor the player with higher bargaining power (Rustichini & Villeval, 2014). Similarly, changes in power influence what is perceived as a fair division: Responders in ultimatum games are judged to deserve more when they have greater power (Mallucci et al., 2019). Third, fairness perceptions of monetary divisions are biased by self-serving tendencies for advantaged players (Amasino et al., 2023; Babcock & Loewenstein, 1997; Rode & Le Menestrel, 2011) and, sometimes, by self-undermining tendencies for disadvantaged players (Lois & Riedl, 2022).

While suggestive, these findings leave several questions unanswered. Does bargaining power influence moral judgment beyond the abstract context of fairness judgments in structured economic games? Does it affect the moral judgments of neutral third parties, or merely bias the judgments of affected parties? And, is there an effect of *stakes* and *available alternatives* on moral judgment when these factors are isolated from other forms of asymmetry, such as the structured roles of "proposer" and "responder"?

To answer these questions, we investigate the effect of manipulations of bargaining power via *stakes* and *available alternatives* on moral judgment in five preregistered controlled experiments ($n = 3,025$) using social scenarios depicting real-life social interactions. In the spirit of the above example involving a graduate student and her advisor, in each scenario, two characters can perform a mutually beneficial but unpleasant action. Characters only differ in terms of their relative bargaining power. The first four studies are based on the same seven social scenarios. Here, bargaining power is mainly manipulated via *stakes*—the *value* of the outcome at stake for each party. Study 1 focuses on the moral appropriateness of asking the other party to perform the unpleasant action. Study 2 focuses on the moral appropriateness of explicitly refusing to perform the unpleasant action. Study 3 focuses on the role played by the magnitude of the difference in bargaining power between both parties. Study 4 extends the results of Study 2 to four additional measures of moral judgment (judgments of evaluation, wrongness, blame, and character). Study 5 is based on five additional scenarios in which bargaining power is instead manipulated via the *alternatives available* to the parties. Study 5a focuses on moral judgments about asking the other party to perform the unpleasant action (as in Study 1). Study 5b focuses on moral judgments about explicitly refusing to perform the unpleasant action (as in Study 2). As in Study 4, we collect five measures of moral judgment in Study 5.

We hypothesize that the same request to perform a mutually beneficial but unpleasant action will be perceived as more morally appropriate when it is made by the party with higher bargaining power than when it is made by the party with lower bargaining power (Studies 1 and 5a). Similarly, we predict that refusing to perform the unpleasant action will be perceived as more morally appropriate when the refusal is expressed by the character with higher bargaining power than when it is expressed by the character with lower bargaining power (Studies 2, 4, and 5b). Finally, we hypothesize that this refusal by the party with lower bargaining power will be perceived as more morally appropriate when their absolute bargaining position is stronger as opposed to weaker (Study 3).

Transparency and Openness

We report how we determined sample sizes, all data exclusions (if any), all manipulations, and all measures in the studies. The design, sample size, hypotheses, and analysis plan of all five studies were preregistered. Study materials, preregistration forms, raw data, and analysis scripts for all studies are available on an Open Science Framework (OSF) repository (https://osf.io/9k4gn). This research was conducted under the approval of the institutional review board at Harvard University. All analyses were conducted using R statistical software (R Core Team, 2024). Sample sizes were determined using power analyses as described in the preregistration form of each study. For details, see Supplemental Material.

Study 1

Method

Participants

The preregistration site is OSF (https://osf.io/gdpw7). Four hundred two U.S.-based participants were recruited from Prolific, and 16 participants were excluded from analyses for failing an attention check, as preregistered (final sample age range = 19–85 years; $M = 44.7$; $SD = 13.7$; 195 women; 187 men; four others). We used the following filtering criteria: Participants located in the United States; balanced sample (male and female participants); first language is English; approval rate: 98–100; exclude participants from previous related studies; and rate limiting (which gives priority access to the study to less experienced participants) was deactivated. Participants were paid $0.70 for 5 min of their time ($8.40/hr). As described in the preregistration form, this sample size was calculated to achieve adequate power ($\beta = 0.8$ with $\alpha = .05$) for a comparison (two tailed) between two independent means (independent groups of equal size) to detect an effect size $d = 0.3$ taking into account typical exclusion rates (based on attention checks) for online samples. Power analyses were conducted using G*Power (Faul et al., 2009) and description of power analyses for subsequent studies are available on each study's preregistration form.

Materials

Participants were presented with seven vignettes (an additional vignette was also used to check attention) constructed according to the following structure: Two characters can perform a mutually beneficial but unpleasant action, one of them has higher bargaining power, the other has worse bargaining power. The bargaining power asymmetry is mainly manipulated via the *value* attributed to the outcome at stake by each party (well, deal, paintings, canoe, boss, stadium), and, for one vignette, via the parties' *available alternatives* (cab).[3] Participants were asked: How morally appropriate would it be for X to ask Y to do Z? They provided a rating of moral appropriateness by indicating a number (from 0 to 100) using a labeled slider (0 = extremely inappropriate; 25 = somewhat inappropriate; 50 = neither appropriate nor inappropriate; 75 = somewhat appropriate;

---

[3] In an earlier version of this article, we referred to our manipulations as manipulations of "outside options" understood in the game-theoretic sense of the payoff/utility that each party would get in the absence of agreement. Formally, all our vignettes—whether bargaining power is manipulated via *stakes* or *available alternatives*—can be modeled using the same approach (see Supplemental Material): Two parties with asymmetric bargaining power due to asymmetric disagreement payoffs/utilities ("outside options") can reach a mutually beneficial agreement if one of them incurs a larger cost (unpleasant action) than the other. We decided to drop the "outside options" terminology, which connotes "available alternatives" and obscures the difference between "stakes" and "alternatives," to avoid confusion. We discuss this at length in the Supplemental Material.

100 = extremely appropriate). For each vignette, participants were randomly assigned to one of two conditions. In the "high bargaining power" condition, X is the character with higher bargaining power and Y is the character with lower bargaining power. In the "low bargaining power" condition, X is the character with lower bargaining power, and Y is the character with higher bargaining power. The vignettes were identical in both conditions, only the order of the names in the moral judgment question differed between conditions (i.e., X replaced with Y and Y replaced with X).

Vignettes varied in the nature of the action to be performed (e.g., jump in the cold water, wait in line), the nature of the outcome at stake (e.g., valuable item, financial bonus), the relationship between the characters (e.g., colleagues, friends, strangers), and the context of the interaction (e.g., work, leisure, travel). We used common American male first names for all characters to prevent any potential influence on results of other factors (e.g., gender, race, socioeconomic status) unrelated to our hypotheses. All vignettes described real-life social scenarios. Vignettes are provided below.

Well. Frank accidentally dropped his wedding ring down the well. Sam accidentally dropped a cheap watch down the well. To get their items back, one of them has to go down the well and get muddy while the other holds the rope.

Cab. James can get home in about 2 hours with the train. Bob can get home in a bit more than 1 hour with the train. They decide to take a cab with the company's credit card instead. The cab can drop either of them in 30 minutes and the other one 30 minutes after that.

Deal. Matt and Daniel are colleagues of equal rank working on an important deal. If the deal goes through, Matt will get a $100,000 bonus as he is the one who generated the lead. Daniel wants the deal to go through but he will not get a bonus. Closing the deal requires one of them to immediately take a cab to the airport to get the client's signature before the client's flight takes off. Meanwhile the other simply needs to go upstairs to inform the boss who manages the client relationship.

Paintings. Will and Joe have both been painting outdoors. Will has been working on his painting for months, Joe has been working on his painting for half an hour. They left their paintings up and went to get lunch. It starts to rain. One of them needs to sprint as fast as possible to save the paintings before they are ruined while the other takes shelter from the rain.

Canoe. Michael and David are sharing a canoe. The only items inside the canoe are Michael's very expensive camera equipment and David's lunch bag. The canoe begins to rock after descending the rapids and one of them needs to jump in the cold water immediately to save it from tipping over.

Boss. Tom and Mark prepare a birthday party at the office for their boss. Tom really needs the party to go well because he plans to ask for a promotion soon. Mark is starting a new job at a different company next month anyway. One of them has to pick up the cake from the bakery down the street. Meanwhile the other needs to spend an hour in traffic to collect the present in the city center.

Stadium. Paul and Steve are buying tickets for the football game. Paul has a coupon worth $90 and can get a ticket for $10. Steve will have to pay the full price of $100. One of them has to stand in line at the boutique for one hour to buy the tickets while the other eats an ice cream in the sun.

Attention Check. Chris and Andy are at the restaurant and are almost done with their dinner. This is an attention check. Please ignore the next question and click on number 94 on the slider below

to demonstrate that you are paying attention. One of them needs to pay the bill while the other calls the hotel to cancel their reservation.

In Well, Frank's bargaining power was lower because, in the absence of agreement, Frank would lose his wedding ring (high value), but Sam would only lose his cheap watch (low value). In Cab, James' bargaining power was lower because his best alternative was to get home in 2 hours with the train (worse alternative) while Bob's best alternative was to get home in 1 hour (better alternative). In Deal, Matt's bargaining power was lower because, in the absence of agreement, in addition to the deal not going through, Matt would lose a $100,000 bonus (high value). Daniel, however, did not expect a bonus (low value). In Paintings, Will's bargaining power was lower because, in the absence of agreement, he would lose a painting he had been working on for months (high value). Joe would only lose a painting he had been working on for half an hour (low value). In Canoe, Michael's bargaining power was lower because, in the absence of agreement, he would lose his very expensive camera equipment (high value). David would only lose his lunch bag (low value). In Boss, Tom's bargaining power was lower because, in the absence of agreement, in addition to a ruined party for his boss, his chances of getting a promotion would be lower (high value). This was not the case for Mark, who is about to leave his job for another one (low value). In Stadium, Paul's bargaining power was lower because, in the absence of agreement, he would both miss the game and lose a coupon worth $90 (high value). Steve would simply miss the game (low value).

Procedure

The study was administered using Qualtrics. Participants were invited to participate in a study about social judgment and decision making. All participants read and completed a consent form. They entered their Prolific ID and read the instructions. Then they were presented with seven vignettes (one per page) and an attention check in a random order. For each vignette, participants were randomly assigned to one of two conditions and provided a judgment of moral appropriateness before proceeding to the next vignette. Next, participants could provide their age, gender, and any comments (in a text box) they had about the study. Finally, all participants were presented with a debriefing form.

Following our preregistration, for each vignette, we performed a linear regression with condition as the independent variable. Then, we collapsed moral judgments for all vignettes and performed a mixed-effects linear regression with condition as our independent variable and random intercepts per participant and per vignette.

We hypothesized that the same request (one character asking another character to perform a mutually beneficial but unpleasant action) would be perceived as more morally acceptable when the character making the request has more bargaining power than when the character making the request has less bargaining power.

Results

Results of our preregistered statistical tests supported our hypothesis. The same request to perform a mutually beneficial but unpleasant action was perceived as more morally acceptable when the request was made by the character with higher bargaining power than when it was made by the character with lower bargaining power for all vignettes, see Figure 1 and Table 1. We also found support for our hypothesis after combining judgments for all vignettes and

Figure 1
Study 1 Results
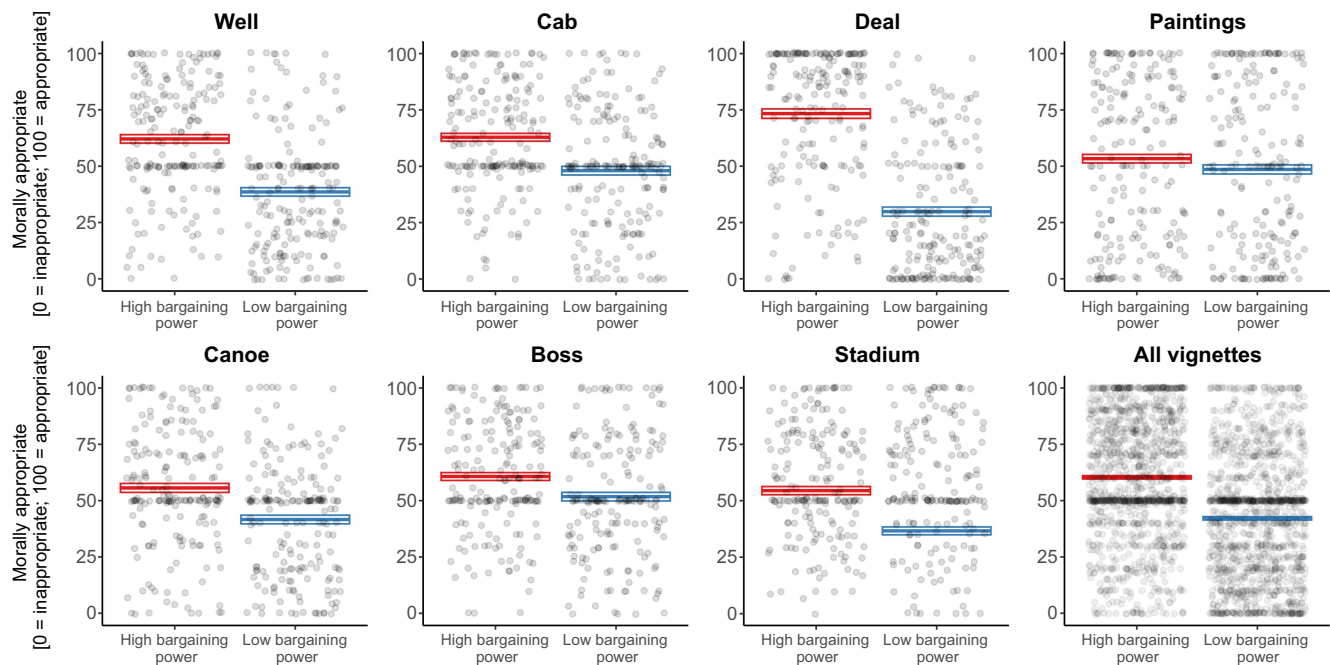
Note. The same request to perform a mutually beneficial but unpleasant action was perceived as more morally appropriate when the request was made by the character with higher bargaining power than when it was made by the character with lower bargaining power. Central bars represent the mean and error bars are standard errors of the mean. See the online article for the color version of this figure.

performing a mixed-effects linear regression with random intercepts per participant and vignette ($b = -17.9$, 95% CI $[-19.9, -16.0]$; $p < .001$; estimated marginal means high: $M = 60.2$; low: $M = 42.3$; Cohen's $d = 0.67$, 95% CI $[0.60, 0.75]$). For detailed results, see Supplemental Material.

Overall, this study suggests that, because of his higher bargaining power, the advantaged party benefits from more moral leeway and can rightfully ask the disadvantaged party to perform the unpleasant action.

Study 2

In Study 1, we found that advantaged parties are considered to have more moral leeway to make polite and nonconfrontational

Table 1
Study 1 Results

Vignette	High BP	Low BP	p	Cohen's d	95% CI
Well	62.1	38.6	<.001	0.91	[0.70, 1.12]
Cab	62.9	48.1	<.001	0.59	[0.38, 0.79]
Deal	73.3	29.8	<.001	1.52	[1.30, 1.75]
Paintings	53.3	48.6	.085	0.18	[−0.02, 0.38]
Canoe	55.6	41.6	<.001	0.54	[0.34, 0.74]
Boss	60.7	51.8	<.001	0.35	[0.15, 0.56]
Stadium	54.5	36.6	<.001	0.72	[0.51, 0.93]
All vignettes	60.3	42.1	<.001	0.67	[0.60, 0.75]

Note. Mean moral appropriateness (0–100) for each condition (BP: high vs. low). *P* values from linear regressions (for each vignette) and from a mixed-effects linear regression with random intercepts per participant and vignette (all vignettes combined). BP = bargaining power; CI = confidence interval.

requests that less advantaged parties take on greater effort for mutual benefit. Study 2 asks whether we observe the same effect for a much more assertive behavior: Refusing to perform the more effortful behavior, implicitly insisting that one's partner must do so instead.

Method

Participants

The preregistration site is OSF (https://osf.io/qm5gb). Four hundred U.S.-based participants were recruited from Prolific, and 34 participants were excluded from analyses for failing an attention check, as preregistered (final sample age range = 19–75 years; $M = 42.6$; $SD = 13.9$; 183 women; 174 men; nine others). We used the same filtering criteria as in Study 1. Participants were paid $0.70 for 5 min of their time ($8.40/hr). The sample size was determined using power analysis as described in the preregistration form.

Procedure

We used the same materials and procedure as in Study 1. For this study, for each vignette, participants were asked: "X explicitly refuses to do Z. As a result, Y does Z. How morally appropriate was it for X to refuse to do Z?" (as in Study 1, Z refers to the unpleasant action: e.g., waiting in line, jumping in the cold water, going down the well). They provided a rating of moral appropriateness by indicating a number (from 0 to 100) using a labeled slider (0 = extremely inappropriate; 25 = somewhat inappropriate; 50 = neither appropriate nor inappropriate; 75 = somewhat appropriate; 100 = extremely

appropriate). For each vignette, participants were randomly assigned to one of two conditions. In the "high bargaining power" condition, X was the character with higher bargaining power, and Y was the character with lower bargaining power. In the "low bargaining power" condition, X was the character with lower bargaining power, and Y was the character with higher bargaining power. The vignettes were identical in both conditions, only the order of the names in the moral judgment question differed between conditions (i.e., X replaced with Y and Y replaced with X).

As preregistered, for each vignette, we performed a linear regression with condition as the independent variable. Then, we collapsed moral judgments for all vignettes and performed a mixed-effects linear regression with condition as our independent variable and random intercepts per participant and per vignette.

We hypothesized that the same refusal (one character explicitly refusing to perform a mutually beneficial but unpleasant action, leading the other character to perform it instead) would be perceived as more morally acceptable when the character expressing the refusal had more bargaining power than when the character expressing the refusal had less bargaining power.

Results

Results of our preregistered statistical tests supported our hypothesis. The same refusal to perform a mutually beneficial but unpleasant action was perceived as more morally acceptable when the refusal was expressed by the character with higher bargaining power
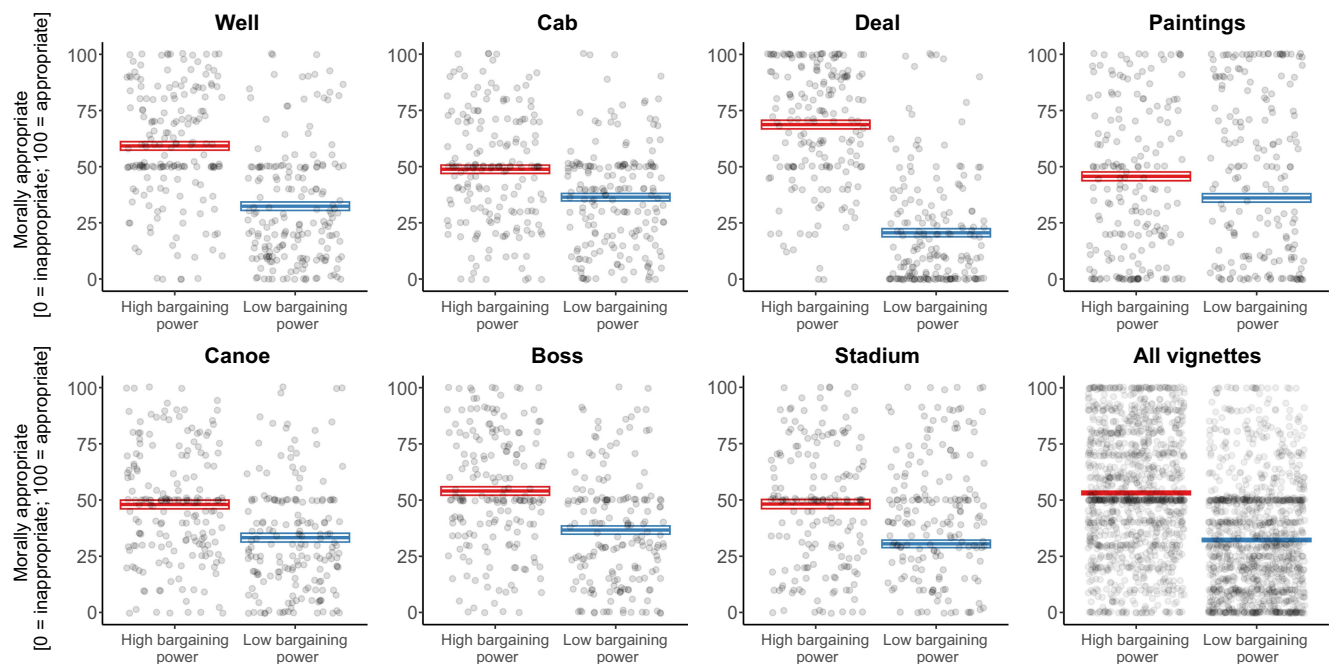
than when it was expressed by the character with lower bargaining power for all vignettes, see Figure 2 and Table 2. We also found support for our hypothesis after combining judgments for all vignettes and performing a mixed-effects linear regression with random intercepts per participant and vignette ($b = -21.2$, 95% CI [-23.0, -19.4]; $p < .001$; estimated marginal means high: $M = 53.3$; low: $M = 32.1$; Cohen's $d = 0.81$, 95% CI [0.73, 0.89]). For detailed results, see Supplemental Material.

Thus, we find that explicitly refusing to perform the unpleasant action—a resolutely confrontational move—is usually appropriate for the party with higher bargaining power but not for the disadvantaged party, which is again held to stricter moral standards.

Study 3

Studies 1 and 2 show that it is considered more morally appropriate for the advantaged party in an interaction to suggest, or insist, that the disadvantaged party take on greater effort for their mutual benefit. In Study 3, we ask whether the magnitude of this effect is related to the magnitude of the relative advantage or disadvantage (i.e., the degree to which the parties diverge in terms of bargaining power). According to the standard bargaining theories, greater asymmetries should lead to greater disparities in the negotiated solution. However, our measure is not the negotiated solution itself, but the moral appropriateness assigned to it. It is possible that people may judge it broadly appropriate for advantaged parties to exert less effort without regard to the size of the relative advantage.

Figure 2

Study 2 Results

Note. The same refusal to perform a mutually beneficial but unpleasant action was perceived as more morally appropriate when the refusal was expressed by the character with higher bargaining power than when it was expressed by the character with lower bargaining power. Central bars represent the mean and error bars are standard errors of the mean. See the online article for the color version of this figure.

Table 2

Study 2 Results

Vignette	High BP	Low BP	*p*	Cohen's *d*	95% CI
Well	59.2	32.4	<.001	1.05	[0.83, 1.27]
Cab	48.8	36.4	<.001	0.52	[0.31, 0.73]
Deal	68.8	20.6	<.001	1.92	[1.67, 2.16]
Paintings	45.7	36.1	<.001	0.37	[0.16, 0.58]
Canoe	48.0	33.3	<.001	0.61	[0.40, 0.82]
Boss	54.0	36.7	<.001	0.69	[0.47, 0.90]
Stadium	48.1	30.5	<.001	0.70	[0.49, 0.91]
All vignettes	53.2	32.2	<.001	0.81	[0.73, 0.89]

Note. Mean moral appropriateness (0–100) for each condition (BP: high vs. low). *P* values from linear regressions (for each vignette) and from a mixed-effects linear regression with random intercepts per participant and vignette (all vignettes combined). BP = bargaining power; CI = confidence interval.

Method

Participants

The preregistration site is OSF (https://osf.io/9cz5u). Eight hundred fourteen U.S.-based participants were recruited from Prolific, and 60 participants were excluded from analyses for failing an attention check, as preregistered (final sample age range = 19–77 years; *M* = 41.1; *SD* = 14.0; 372 women; 361 men; 21 others). We used the same filtering criteria as in previous studies. Participants were paid $0.70 for 5 min of their time ($8.40/hr). The sample size was determined using power analysis as described in the preregistration form.

Procedure

We used the same vignettes as in Study 1 and Study 2. Participants were asked: "X explicitly refuses to do Z. As a result, Y does Z. How morally appropriate was it for X to refuse to do Z?" They provided a rating of moral appropriateness by indicating a number (from 0 to 100) using a labeled slider (0 = extremely inappropriate; 25 = somewhat inappropriate; 50 = neither appropriate nor inappropriate; 75 = somewhat appropriate; 100 = extremely appropriate). For each vignette, participants were randomly assigned to one of three conditions. In each condition, X had lower bargaining power than Y. We manipulated the magnitude of the asymmetry in bargaining power between X and Y by varying X's *stakes* in the agreement, and thus the strength of his bargaining position. The stakes for X were lower in the "low stakes" condition (stronger bargaining position) than in the "medium stakes" condition (medium bargaining position) and the "high stakes" condition (weaker bargaining position).[4] The vignettes were identical in all conditions, only X's stakes in the agreement varied.

Well. Frank accidentally dropped his bracelet, a gift from his wife/gold chain, a gift from his wife/wedding ring down the well.

Cab. James can get home in about 1.5 hours/2 hours/3 hours with the train.

Deal. Matt will get a $1,000/$10,000/$100,000 bonus as he is the one who generated the lead.

Paintings. Will has been working on his painting for a month/3 months/12 months.

Canoe. The only items inside the canoe are Michael's camera equipment worth $500/$3,000/$10,000 and David's lunch bag.

Boss. Tom really needs the party to go well because he plans to ask for a 2%/10%/30% wage increase soon.

Stadium. Paul has a coupon worth $20 and can get a ticket for $80/$50 and can get a ticket for $50/$90 and can get a ticket for $10.

As preregistered, we collapsed moral judgments for all vignettes and performed a mixed-effects linear regression with condition (ordered factor) as our independent variable and random intercepts per participant and per vignette.

We hypothesized that the same refusal (the character with less bargaining power explicitly refusing to perform a mutually beneficial but unpleasant action, leading the other character with more bargaining power to perform it instead) would be perceived as more morally appropriate when the character's bargaining power was stronger as opposed to weaker. We predicted that there would be evidence in support of a linear trend between moral judgments and the strength of the character's bargaining position (moral judgments go down as the bargaining position of the character goes down; conversely, moral judgments go up as the bargaining position of the character goes up).

Results

While the results of our preregistered statistical test (mixed-effects linear regression with condition [ordered factor] as our independent variable and random intercepts per participant and per vignette) were consistent with our predictions, they suggest that, contrary to our preregistered predictions, the influence of the strength of the bargaining position on moral judgments was negligible, see Figure 3 and Table 3 ($b = -1.1$, 95% CI [-2.0, -0.1]; $p = .024$; estimated marginal means high: $M = 32.5$; medium: $M = 31.2$; low: $M = 31.0$; comparison between high and low: Cohen's $d = 0.06$, 95% CI [0.00, 0.13]).

We confirmed this by running exploratory linear regressions with condition (ordered factor) as our independent variable for each vignette, see Table 3. The difference between conditions was negligible and not significant for all vignettes which suggests that, when the character with lower bargaining power explicitly refuses to perform the mutually advantageous but unpleasant action, the strength of their bargaining position did not influence the moral appropriateness of the refusal. For detailed results, see Supplemental Material.

Overall, the relative position in the potential negotiation, as opposed to the magnitude of the bargaining power asymmetry between the parties, seems to drive perceived differences in moral appropriateness.

Study 4

Thus far our studies have only used one measure of moral judgment, moral appropriateness, which belongs to the category of "norm" judgments (Malle, 2021). In Study 4, we ask whether the observed effect also holds for other types of measures of moral judgment, including judgments of evaluation, wrongness, blame, and character.

---

[4] For the cab vignette, we varied the value of the low bargaining power party's alternative, which could be high (stronger bargaining position), medium (intermediate bargaining position), or low (weaker bargaining position).

Figure 3

Study 3 Results
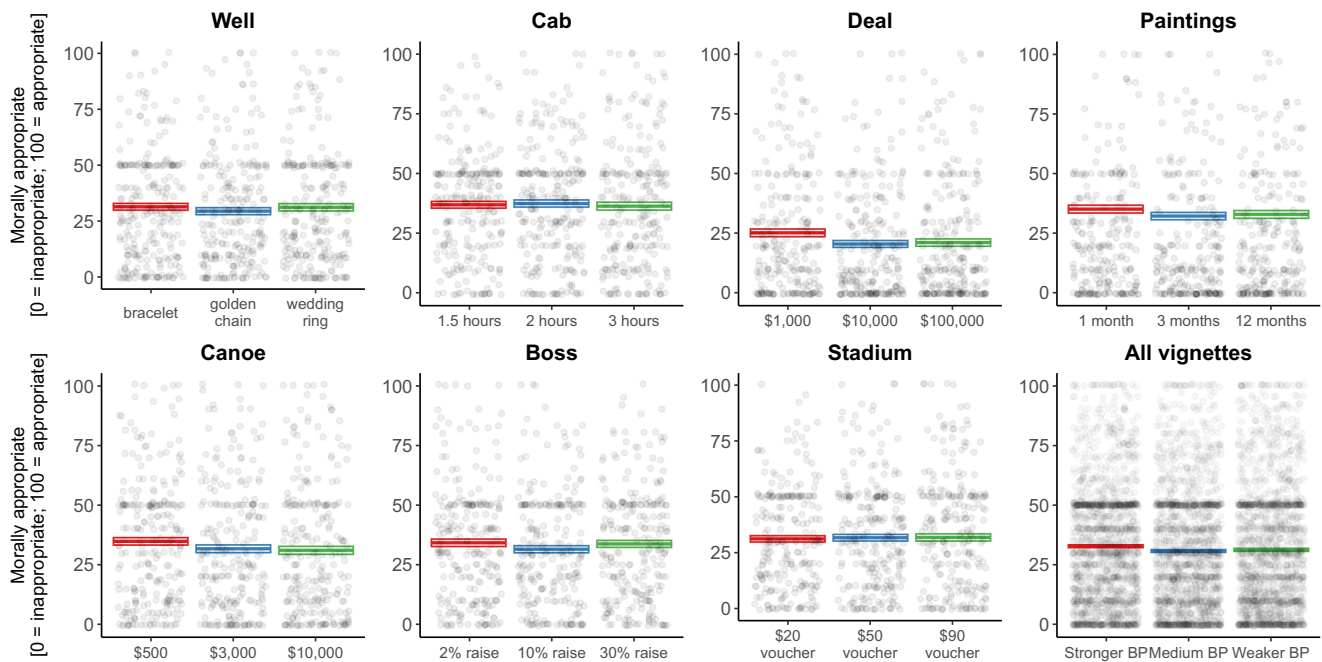
Note. Explicit refusal to perform the mutually beneficial but unpleasant action by the character with lower bargaining power was perceived as morally inappropriate irrespective of the strength of their bargaining position, manipulated via the value of the outcome at stake. The influence of the magnitude of the difference in stakes on moral judgments was negligible and not systematic. Central bars represent the mean and error bars are standard errors of the mean. See the online article for the color version of this figure.

Method

Participants

The preregistration site is OSF (https://osf.io/4gz8n). Four hundred four U.S.-based participants were recruited from Prolific, and 40 participants were excluded from analyses for failing an attention check, as preregistered (final sample age range = 18–82 years; $M = 43.0$; $SD = 13.0$; 185 women; 174 men; five others). We used the same filtering criteria as in the previous studies. Participants were paid $1.40 for 10 min of their time ($8.40/hr). The sample size

was determined using power analysis as described in the preregistration form.

Procedure

We used the same vignettes as in Studies 1–3. As in Studies 2 and 3, for each vignette participants read: "X explicitly refuses to do Z. As a result, Y does Z." Then they were asked to make five moral judgments. Each of the first four measures corresponded to one of the major classes of moral judgment according to Malle (2021):

Table 3

Study 3 Results

Vignette	Stronger BP	Medium BP	Weaker BP	p	Cohen's d	95% CI
Well	31.5	29.6	31.3	.924	0.01	[−0.17, 0.18]
Cab	37.1	37.6	36.5	.785	0.02	[−0.15, 0.20]
Deal	25.4	20.8	21.3	.059	0.16	[−0.01, 0.34]
Paintings	35.3	32.4	33.1	.332	0.09	[−0.09, 0.26]
Canoe	34.8	31.7	31.1	.112	0.14	[−0.03, 0.32]
Boss	34.2	31.5	33.7	.818	0.02	[−0.15, 0.20]
Stadium	31.1	31.6	31.7	.760	−0.03	[−0.20, 0.15]
All vignettes	32.8	30.8	31.3	.024	0.06	[0.00, 0.13]

Note. Mean moral appropriateness (0–100) for each condition (BP: strength of the bargaining position for the character with lower bargaining power: stronger, medium, weaker). P values from linear regressions (for each vignette) and from a mixed-effects linear regression with condition (ordered factor) as the independent variable and random intercepts per participant and per vignette (all vignettes combined). Cohen's d: comparison between stronger BP and weaker BP. BP = bargaining power; CI = confidence interval.

norm, evaluation, wrongness, and blame judgments. We also added a fifth measure of moral character. For each vignette, the first measure (norm) was our usual measure of moral appropriateness—our primary dependent variable—the other four measures were presented in random order. All measures were presented on the same page as the vignette.

Norm. "How morally appropriate was it for X to refuse to do Z?" (labeled slider [0 = extremely inappropriate; 25 = somewhat inappropriate; 50 = neither appropriate nor inappropriate; 75 = somewhat appropriate; 100 = extremely appropriate]).

Evaluation. "How morally good or bad was it for X to refuse to do Z?" (labeled slider [0 = extremely bad; 25 = somewhat bad; 50 = neither good nor bad; 75 = somewhat good; 100 = extremely good]).

Wrongness. "How morally right or wrong was it for X to refuse to do Z?" (labeled slider [0 = extremely wrong; 25 = somewhat wrong; 50 = neither right nor wrong; 75 = somewhat right; 100 = extremely right]).

Blame. "How morally praiseworthy or blameworthy was X for refusing to do Z?" (labeled slider [0 = extremely blameworthy; 25 = somewhat blameworthy; 50 = neither praiseworthy nor blameworthy; 75 = somewhat praiseworthy; 100 = extremely praiseworthy]).

Character. "To what extent does X seem like a morally good or bad person to you?" (labeled slider [0 = extremely bad person; 25 = somewhat bad person; 50 = neither good nor bad person; 75 = somewhat good person; 100 = extremely good person]).

Apart from these additional measures, the procedure was identical to that of Study 2, with participants being presented with all vignettes in random order and randomly allocated to one of two versions of each vignette.

As preregistered, for each measure, we performed a mixed-effects linear regression with condition as our independent variable and random intercepts per participant and vignette. Then, we collapsed moral judgments for all measures and performed a mixed-effects linear regression with condition as our independent variable and random intercepts per participant, vignette, and measure.

We hypothesized that, for each measure, the same refusal (one character explicitly refusing to perform a mutually beneficial but unpleasant action, leading the other character to perform it instead) would be perceived as morally worse when the character expressing the refusal had less bargaining power than when the character expressing the refusal had more bargaining power.

Results

Results of our preregistered statistical tests supported our hypotheses. For all measures, the same refusal to perform a mutually beneficial but unpleasant action was perceived as morally worse when the refusal was expressed by the character with lower bargaining power than when it was expressed by the character with higher bargaining power, see Figure 4 and Table 4. We also found support for our hypothesis after combining judgments for all measures and performing a mixed-effects linear regression with random intercepts per participant, vignette, and measure ($b = -15.0$, 95% CI [-15.6, -14.3]; $p < .001$; estimated marginal means high: $M = 46.4$; low: $M = 31.4$; Cohen's $d = 0.73$, 95% CI [0.69, 0.76]). For detailed results for each vignette, see Supplemental Material.

Thus, the observed asymmetry is not limited to moral appropriateness but also holds for other measures of moral judgment.

Figure 4
Study 4 Results
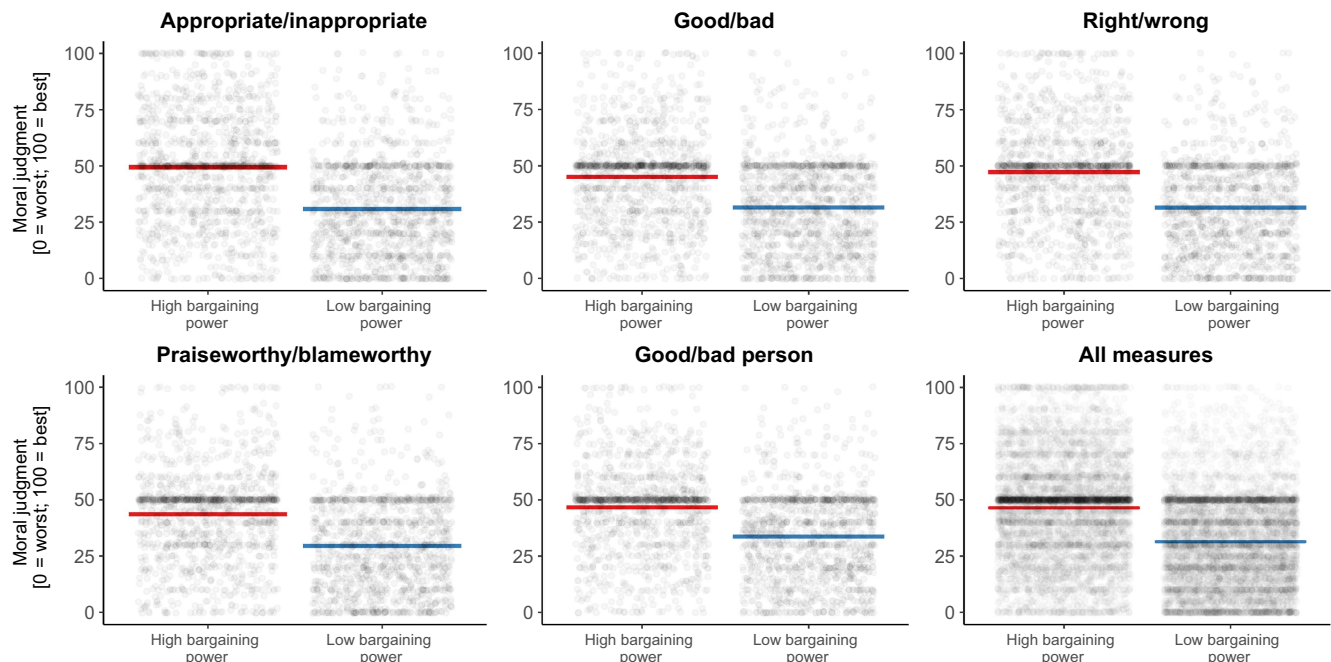
Note. Results are similar for five types of measures of moral judgment. The same refusal to perform a mutually beneficial but unpleasant action was perceived as morally worse when the refusal was expressed by the character with lower bargaining power than when it was expressed by the character with higher bargaining power. Central bars represent the mean and error bars are standard errors of the mean. See the online article for the color version of this figure.

Table 4

Study 4 Results

Measure	High BP	Low BP	p	Cohen's d	95% CI
Appropriate/inappropriate	49.4	30.9	<.001	0.81	[0.73, 0.89]
Good/bad	45.1	31.5	<.001	0.69	[0.61, 0.77]
Right/wrong	47.3	31.5	<.001	0.75	[0.67, 0.83]
Praiseworthy/blameworthy	43.6	29.5	<.001	0.71	[0.63, 0.79]
Good/bad person	46.6	33.6	<.001	0.67	[0.59, 0.75]
All measures	46.4	31.4	<.001	0.73	[0.69, 0.76]

Note. Mean moral judgment (0–100) for each measure (BP: high vs. low). *P* values from mixed-effects linear regressions with random intercepts per participant and vignette (for each measure) and mixed-effects linear regression with random intercepts per participant, vignette, and measure (all measures combined). BP = bargaining power; CI = confidence interval.

Study 5

The vignettes used in Studies 1–4 mainly involved manipulations of bargaining power via the *value* attributed to the outcome at stake by each party. Another way to manipulate bargaining power is to endow one party with (a) better *alternative(s)* than the other in the absence of agreement (e.g., as in the cab vignette where one character can go home faster with the train than the other if they cannot reach an agreement). We construct five new vignettes based on the same previous structure—two parties mutually benefit if one of them performs an unpleasant action—in which the characters differ in their bargaining power based on the *alternatives available* to them as opposed to the *value* they attribute to the outcome at stake. In Study 5a, we ask about the moral appropriateness of asking the other party to perform the unpleasant action (as in Study 1). In Study 5b, we ask about the moral appropriateness of explicitly refusing to perform the unpleasant action (as in Studies 2 and 4).

Method

Participants

Study 5a. The preregistration site is OSF (https://osf.io/ahmc8). Five hundred two U.S.-based participants were recruited from Prolific, and 42 participants were excluded from analyses for failing an attention check, as preregistered (final sample age range = 18–77 years; *M* = 40.0; *SD* = 12.3; 233 women; 218 men; nine others).

Study 5b. The preregistration site is OSF (https://osf.io/sne8m). Five hundred three U.S.-based participants were recruited from Prolific, and 31 participants were excluded from analyses for failing an attention check, as preregistered (final sample age range = 18–78 years; *M* = 39.2; *SD* = 12.7; 243 women; 220 men; nine others).

We used the same filtering criteria as in previous studies. Participants were paid $1.10 for 8 min of their time ($8.25/hr). Sample sizes were determined using power analysis as described in the preregistration form.

Materials

Participants were presented with five vignettes (an additional vignette was also used to check attention), which are provided below. Each vignette follows the same structure: Two parties can mutually benefit if one of them performs an unpleasant action. Parties only

differ in their bargaining power, manipulated via alternatives: One of them has (a) better alternative(s) than the other in the absence of agreement and therefore higher bargaining power.

Partners. Brian and George work in the same industry. To finalize an important and mutually beneficial deal, one of them needs to travel to the other's office, which takes an entire day. But each of them is quite busy these days. They both know that, if they cannot meet up in person shortly, Brian can make a similar deal with another business partner instead. George, on the other hand, has no suitable alternative, and losing the deal would put his company at risk.

Project. Simon and John are taking the same course at university. Students must find a partner for a group project and each group must pick a topic on a first-come first-served basis. They both know that Simon is the best student and that he could easily find another partner if needed. John has been struggling a bit in this class and his other options are more limited. One of them needs to go to the professor's office on the other side of campus to sign up the group on the registration form and confirm their choice of topic to the professor.

Volleyball. Jim and Vince are at a summer camp. Registrations for the camp's beach volleyball tournament are open. They both know that Jim is the best player and that he could easily find another teammate if needed. Because he is less skilled, Vince's alternative options are quite limited. One of them needs to leave the beach early to go to the camp's reception and register the team before the deadline.

Construction. Ray is a plumber and Mark is an electrician. They have an opportunity to join forces for a big project at a new construction site. To secure the deal one of them must go to the client's office to go over some details, which will take a few hours. But each of them is quite busy these days. They both know that Ray is already established in the area and can easily find similar projects with other electricians if needed. Mark, on the other hand, is a newcomer and urgently needs new clients.

Condo. Ronald and Jason live in separate flats in the same condo. The condo's heating system has suddenly stopped working and a snow storm is expected this weekend. One of them needs to deal with the heating maintenance company today, which is tedious and time consuming. They both know that, if needed, Jason can stay at his partner's place across town. Ronald, on the other hand, will have to stay in the condo during the snow storm.

In partners, Brian's bargaining power is higher because his alternative is to make a deal with another business partner (better alternative) while George's bargaining power is lower because he has no suitable alternative (worse alternative). In project, Simon's

bargaining power is higher because he can easily find another partner for the group project (better alternatives) while John's bargaining power is lower because his options are more limited (worse alternatives). In volleyball, Jim's bargaining power is higher because he can easily find another teammate for the tournament (better alternatives) while Vince's bargaining power is lower because his alternatives are more limited (worse alternatives). In construction, Ray's bargaining power is higher because, in the absence of agreement, he can easily find a similar project with other electricians instead (better alternatives) while Mark's other options are worse because he is a newcomer in the area (worse alternatives). In condo, Jason's bargaining power is higher because, in the absence of agreement, he can stay at his partner's place (better alternative) while Ronald will have to stay in the condo during the snowstorm (worse alternative).

Procedure

We used the same procedure as in Study 4. In Study 5a, for each vignette and each measure, participants were asked: "How morally … would it be for X to ask Y to do Z?" (minor variations in wording to match each measure, see Supplemental Material). In Study 5b, for each vignette, participants read: "X explicitly refuses to do Z. As a result, Y does Z." Then they were asked to make five moral judgments. Participants were presented with all vignettes in random order and randomly allocated to one of two versions of each vignette.

As preregistered, for each vignette, we performed a mixed-effects linear regression with condition as our independent variable and

random intercepts per participant and measure. Then, we collapsed moral judgments for all vignettes and performed a mixed-effects linear regression with condition as our independent variable and random intercepts per participant, vignette, and measure.

We hypothesized that, for each vignette, the same request (asking the other party to perform the unpleasant action, Study 5a) or refusal (one character explicitly refusing to perform a mutually beneficial but unpleasant action, leading the other character to perform it instead, Study 5b) would be perceived as morally worse when the character expressing the refusal had less bargaining power than when the character expressing the refusal had more bargaining power.

Results

Overall, results of our preregistered statistical tests provided support for our hypotheses in Study 5. While we only observed small effects in Study 5a, effect sizes were substantially larger in Study 5b.

Study 5a

As predicted, for four vignettes (partners, project, volleyball, construction), the same request to perform a mutually beneficial but unpleasant action was perceived as morally worse when the request was made by the character with lower bargaining power than when it was made by the character with higher bargaining power, see Figure 5 and Table 5. But the difference was not statistically significant for two vignettes (project and construction). And, contrary to our predictions, for one vignette (condo) the same request was perceived as morally

Figure 5
Study 5a Results
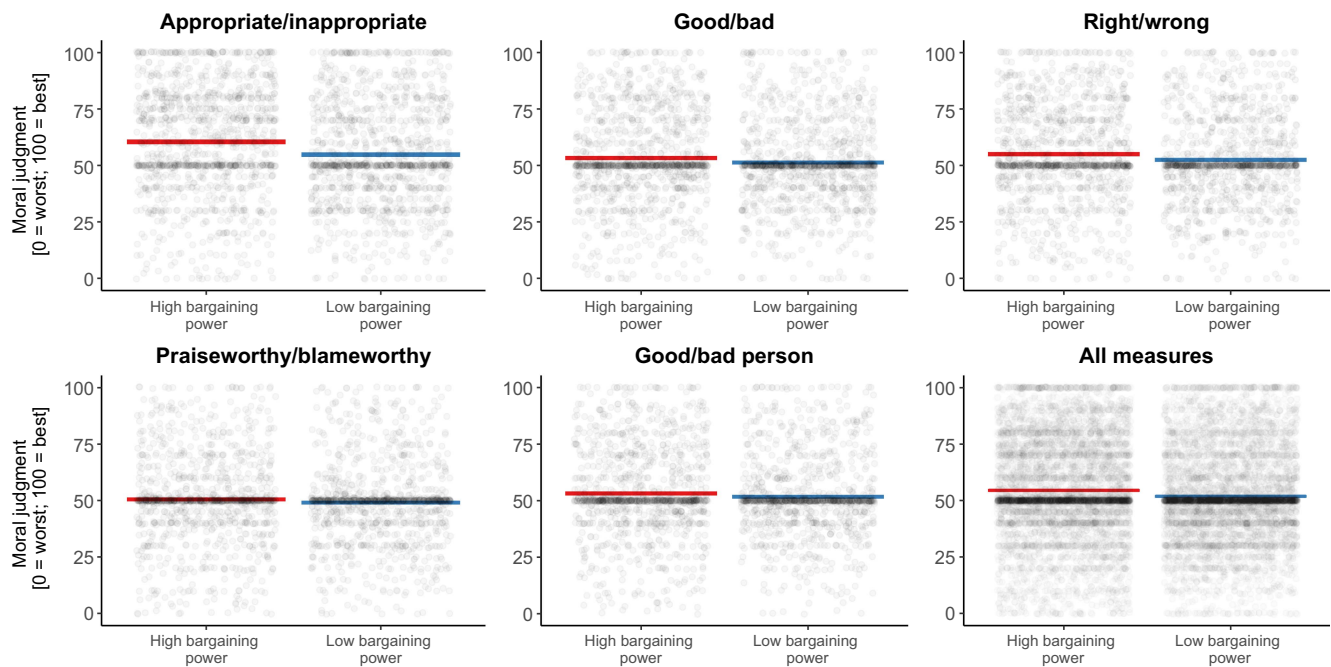
Note. When manipulating bargaining power via available alternatives for each party, the same request to perform a mutually beneficial but unpleasant action was perceived as morally worse (average of all five measures) when it was made by the character with lower bargaining power than when it was made by the character with higher bargaining power. Central bars represent the mean and error bars are standard errors of the mean. See the online article for the color version of this figure.

Table 5
Study 5a Results

Vignette	High BP	Low BP	*p*	Cohen's *d*	95% CI
Partners	55.9	50.2	.00146	0.26	[0.18, 0.34]
Project	54.6	52.3	.148	0.12	[0.04, 0.20]
Volleyball	55.4	50.6	.00194	0.26	[0.18, 0.34]
Construction	57.6	55.0	.126	0.13	[0.05, 0.21]
Condo	48.9	51.3	.129	−0.13	[−0.21, −0.04]
All vignettes	54.5	51.9	<.001	0.13	[0.09, 0.17]

Note. Mean of five moral judgment measures (0–100) for each vignette (BP: high vs. low). *P* values from mixed-effects linear regressions with random intercepts per participant and measure (for each vignette) and mixed-effects linear regression with random intercepts per participant, vignette, and measure (all vignettes combined). BP = bargaining power; CI = confidence interval.

better for the party with lower bargaining power, though this difference was not statistically significant either. We found overall support for our hypothesis after combining judgments for all vignettes and performing a mixed-effects linear regression with random intercepts per participant, vignette, and measure ($b = -2.3$, 95% CI [−2.9, −1.7]; $p < .001$; estimated marginal means high: $M = 54.3$; low: $M = 52.0$; Cohen's $d = 0.13$, 95% CI [0.09, 0.17]). However, the size of the observed effect was small. For detailed results for each measure, see Supplemental Material.

Study 5b

The same refusal to perform a mutually beneficial but unpleasant action was perceived as morally worse when the refusal was expressed by the character with lower bargaining power than when it was expressed by the character with higher bargaining power, see Figure 6 and Table 6. For one vignette (condo), the difference is in the predicted direction but not statistically significant. We also found support for our hypothesis after combining judgments for all vignettes and performing a mixed-effects linear regression with random intercepts per participant, vignette, and measure ($b = -7.0$, 95% CI [−7.6, −6.4]; $p < .001$; estimated marginal means high: $M = 39.9$; low: $M = 32.9$; Cohen's $d = 0.44$, 95% CI [0.40, 0.48]). For detailed results for each measure, see Supplemental Material.

Overall, we observed the same asymmetry described in Studies 1–4 after manipulating bargaining power via *available alternatives* as opposed to the *value* at stake for each party. While the observed effect is small in Study 5a (request), it is substantially larger in Study 5b (refusal).[5] We suggest two potential explanations. First, simply asking the other party to perform the unpleasant action is overall less morally problematic than expressing explicit refusal—as reflected by the higher mean moral judgments (high BP: 54.5, low BP: 51.9) in Study 5a than in Study 5b (high BP: 40.9, low BP: 31.9). It is plausible that the effect of bargaining power (perhaps specifically when it is manipulated via the quality of available alternatives) on moral judgment interacts with the overall acceptability of the underlying behavior, which could explain why we observe smaller effects in Study 5a. Second, it seems that a reasonable inference participants can make when reading the condo vignette is that there is something (left unsaid) that prevents the low bargaining power party from dealing with the heating maintenance company (otherwise why take the risk of having no heating during the snowstorm?) and

that he is simply asking for urgent help (whereas the high bargaining power party can easily be perceived as lazy or uncooperative). This may explain why we do not find the predicted effect for that vignette, which contributes to driving the size of the overall effect down.

Constraints on Generality

We note the following limitations to the generalizability of the present research. Our studies only employ 12 vignettes which clearly do not exhaust the range of possible types of relationships, social interactions, and contexts in which bargaining power plays an important role. Similarly, for reasons of experimental control, interactions are limited to characters with American male first names. Thus, future research should investigate to what extent the relationship between bargaining power and moral judgment depends on the type of social relationship (e.g., hierarchical, romantic, familial), the context (e.g., organizational, political), and demographic factors such as gender, race, or socioeconomic status. We also note that there is variability in the size of the effect depending on the specifics of each vignette, the measures of moral judgment used, the type of behavior being evaluated (request or explicit refusal), and the type of bargaining power manipulation (stakes or available alternatives). In addition, several of our vignettes showing relatively larger effect sizes (e.g., deal, partners, construction) involve contexts of business negotiations. It is thus possible that moral norms in such restricted contexts entail specific inferences about bargaining power, which future research should investigate in more depth.[6] Next, our studies are limited to hypothetical moral judgments with no real stakes for participants. Past research has found that hypothetical judgments often differ from actual actions (e.g., incentivized decisions; FeldmanHall et al., 2012) and do not necessarily reflect real-life behavior (Bostyn et al., 2018). Thus, future studies should investigate the role played by bargaining power on incentivized moral decisions or in real-word settings. In addition, our studies are limited to dyads with one-shot opportunities for mutually beneficial interaction, whereas many interactions occur in contexts of repeated interactions over long time periods and are primarily governed by other factors like reciprocity (Trivers, 1971). Finally, our studies are limited to online samples of U.S. participants which are not representative of the American or global population and are likely to be substantially Western, educated, industrialized, rich, and democratic (Henrich et al., 2010). As such, the extent to which our findings will generalize to other countries and cultures is unclear.

General Discussion

Does bargaining power influence moral judgment? We constructed different scenarios in which two characters can achieve a mutually beneficial outcome if one of them performs an unpleasant action. The characters only differed in their bargaining power. In a first set of scenarios (Studies 1–4), the asymmetry in bargaining power was mainly manipulated via the *stakes* that each party has in the potential agreement. In a second set of vignettes (Study 5), characters differed in their bargaining power due to the quality of the *alternatives available* to them in the absence of agreement. In each study (except Study 3),

---

[5] Importantly, the cab vignette (from the first set of vignettes), in which bargaining power is also manipulated via alternatives available to each party, shows a sizable effect in the predicted direction for both request (Study 1) and refusal (Studies 2 and 4).

[6] We thank an anonymous reviewer for pointing this out.

Figure 6

Study 5b Results
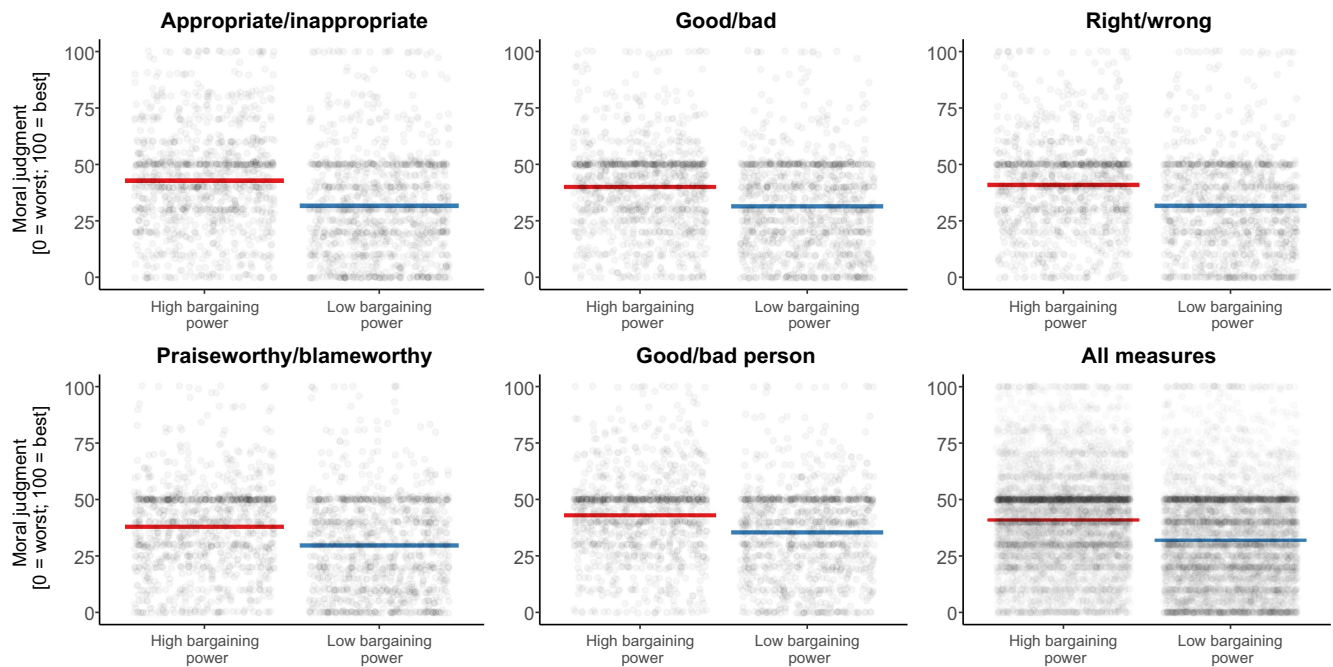
Note. When manipulating bargaining power via alternatives available to each party, the same refusal to perform a mutually beneficial but unpleasant action was perceived as morally worse (average of all five measures) when the refusal was expressed by the character with lower bargaining power than when it was expressed by the character with higher bargaining power. Central bars represent the mean and error bars are standard errors of the mean. See the online article for the color version of this figure.

for each scenario, we randomized participants into two conditions and asked them to rate how morally appropriate it would be for one of the characters to ask the other to perform the unpleasant action. We found that the same behaviors (asking the other party to perform the unpleasant action [Studies 1 and 5a], explicitly refusing to do it [Studies 2, 4, and 5b]) were perceived as more morally appropriate for the party with higher bargaining power than for the party with lower bargaining power. In Studies 4 and 5, we extend our results to five measures of moral judgment (norm, evaluation, wrongness, blame, and character), and find the predicted effect with comparable effect sizes for all measures. Finally, contrary to our predictions, results of Study 3 suggest that it is the relative bargaining power in the

Table 6

Study 5b Results

Vignette	High BP	Low BP	*p*	Cohen's *d*	95% CI
Partners	43.8	29.7	<.001	0.68	[0.60, 0.77]
Project	42.4	32.9	<.001	0.47	[0.39, 0.55]
Volleyball	42.8	35.8	<.001	0.37	[0.29, 0.45]
Construction	44.5	32.7	<.001	0.57	[0.49, 0.65]
Condo	31.1	28.6	.133	0.13	[0.05, 0.21]
All vignettes	40.9	31.9	<.001	0.44	[0.40, 0.48]

Note. Mean of five moral judgment measures (0–100) for each vignette (BP: high vs. low). *P* values from mixed-effects linear regressions with random intercepts per participant and measure (for each vignette) and mixed-effects linear regression with random intercepts per participant, vignette, and measure (all vignettes combined). BP = bargaining power; CI = confidence interval.

interaction (having a higher or lower bargaining power than the other party) as opposed to the absolute difference in bargaining power that is responsible for the observed differences in moral judgments. Taken together, our findings suggest that bargaining power asymmetries can substantially shape the perceived moral appropriateness of certain behaviors, to the benefit of the party with higher bargaining power, and to the detriment of the disadvantaged party.

This pattern of results is broadly consistent with contractualist views of moral cognition, according to which bargaining power guides which tacit agreements are most likely to be followed which, in turn, influences what we perceive as morally appropriate (André et al., 2023; Levine, Chater, et al., 2024). It can also help us to distinguish between different varieties of cognitive contractualism that have been suggested in the literature. In particular, these findings give credence to views that insist on the crucial role played by negotiation-based reasoning processes like virtual bargaining (Chater et al., 2022), which propose that we behave according to the agreements that people would come to given their actual positions in the world (see also: Le Pargneux et al., 2024; Levine, Kleiman-Weiner, et al., 2024). These results are harder to explain for contractualist accounts that emphasize egalitarian bargaining solutions, for instance because bargainers are placed behind a veil of ignorance and therefore have equal bargaining power (Binmore, 2005), or because bargainers start from an initial position that prevents them from taking advantage of their bargaining power (Gauthier, 1986). Rather, from an egalitarian viewpoint, disadvantaged parties (with lower bargaining power) and advantaged parties (with higher bargaining power) should have equal moral leeway, especially in the context of interactions with

opportunities for mutual advantage. While it is possible that such egalitarian views guided the judgments of some of our participants, they cannot account for the total pattern of behavior across the population. Rather, at least some participants' judgments track what we observe in historically unequal distributions of resources, the preponderance of unfair norms in human societies (Guala, 2013), and observed preferences for unequal societies (Starmans et al., 2017). Our results are also in line with previously reported results in the experimental economics literature, where the fairness judgments of participants in a game typically favor more resources for the advantaged player (Lois & Riedl, 2022; Mallucci et al., 2019; Rode & Le Menestrel, 2011; Rustichini & Villeval, 2014).

Might our results be explained by theories that do not appeal to bargaining or negotiation at all? Two such possibilities stand out as especially deserving of further attention, and each will require additional research to conclusively address. First, people may consider the most equitable distribution of *effort* to track the distribution of anticipated *benefit*. If benefit is defined relative to each party's disagreement outcome, then the disadvantaged party stands to realize a greater benefit, and thus (participants might reason) it is fair that they should pay greater effort costs. Figuring out "who has more at stake" seems to us a particularly plausible heuristic that many participants may be consciously adopting when responding to our vignettes.[7] As a heuristic, it is deeply intertwined with bargaining principles, but avoids the necessity of consciously or explicitly reasoning about bargaining power in the way that economists would, a process that is likely to require training and effortful multistep deliberation. For example, a research mentor could reason "My student has more to gain by publishing this paper quickly, so it's only fair that they handle the tedious submission process." Here, in effect, the mentor takes advantage of their better bargaining position to conclude that the PhD student *should* handle the submission.

As such, this first interpretation stands out as a potential psychological model of participants' judgments in at least some cases and does not make explicit reference to bargaining or negotiation at the "proximate" level of psychological mechanism. Further research on this point is warranted. Turning to the "ultimate" level of adaptive design, one reason people might hold such a theory of fair effort allocation (i.e., one grounded in an analysis of bargaining power) is because it provides a cognitively efficient heuristic that closely tracks the outcomes of negotiations predicted by game theory. These intuitions may have been shaped by cultural evolution, for instance, or may be an efficient summary of their own prior reasoning or experience. In other words, it may be a "cognitive resource-rational" approximation of contractualist reasoning. Along these lines, André et al. (2023) and Levine, Chater, et al. (2024) have argued that many of our intuitions about fair distributions of resources, such as inequity aversion, are ultimately established by the logic of bargaining and agreement.

The second possibility has a similar structure but posits a utilitarian method of moral evaluation (Crockett, 2013; Cushman, 2013; Greene, 2014). For instance, if one makes certain assumptions about diminishing marginal utility in our vignettes, aggregate utility might be maximized by having the biggest "beneficiary" bear the highest costs. Similar outcomes might be obtained through assumptions about reference-dependent hedonic responses (i.e., one's positive affect for a mutually beneficial solution being defined against the baseline of one's disagreement outcome). Although accounts of this kind cannot be definitively ruled out based on the data we present

here, several considerations lead us to regard them as unlikely. First, they posit that people reason about relatively sophisticated auxiliary features of others' utilities, and integrate these into their moral judgments. Second, they leave the correspondence between participants' moral judgments and classic theorems of bargaining and negotiation as a "suspicious coincidence." Third, they do not provide a straightforward account for the fairness judgments of participants in behavioral economic games where bargaining power is adjusted through the manipulation of structural roles (e.g., proposer, responder) rather than the players' *stakes in the agreement* or *available alternatives*.

Our findings have important potential implications for our understanding of the emergence and persistence of inequality. If some of our moral intuitions are contractualist and we have a tendency, in specific contexts (e.g., opportunities for mutual advantage), to be more lenient toward those in better bargaining positions (often already better off) and morally stricter with those in worse bargaining positions (often already worse off or disadvantaged), then our moral cognition may facilitate the emergence and persistence of inegalitarian, unequal, and perhaps unfair norms and outcomes. For instance, in partners, the party who is already better off (Brian, who has more business partners) is able to reap the same benefits from the deal as the party who is initially worse off (George) and people view it as appropriate that he should exert *less* effort to do so, further exacerbating the initial inequality in a "rich get richer" dynamic. In this way, our studies complement game theoretic simulations showing how inequality and unfairness can emerge, evolve, and persist under minimal conditions and even in the absence of intended prejudice or discrimination (Bruner & O'Connor, 2016; O'Connor, 2019). Combating inequalities may require going against some of our moral intuitions and a natural inertia that favors those who are already better off.

Last, three limitations of our studies raise intriguing questions for future research. First, our studies do not investigate how bargaining power affects moral judgments when the cost of performing the unpleasant task exceeds its expected benefits, especially for the disadvantaged party. Using one's power (e.g., mafia boss) for extortion or exploitation purposes seems straightforwardly morally wrong, suggesting that the effect of bargaining power on moral judgment is context-dependent and hinges on whether the interaction is expected to be mutually beneficial or to lead to a net loss for one party. Second, our studies are limited to moral judgments made from a third-party perspective, but past research using economic games has shown evidence of self-serving tendencies (Babcock & Loewenstein, 1997) in fairness perceptions. It is therefore interesting to ask if such self-serving concerns also bias moral judgments by each involved party in more ecologically valid social scenarios like ours, and if this effect is symmetric. Third, our vignettes do not exhaust the range of social-relational settings that people find themselves in, and future research should explore the extent to which our results extend beyond market-type interactions to settings of communal sharing, authority ranking, or equality matching (Rai & Fiske, 2011).

---

[7] "I should act when I have more at stake in the outcome" is a plausible candidate for a principle that could be impartially justified, and that relevant stakeholders could not reasonably reject. Under this interpretation it might be possible for more "egalitarian" contractualist theories (e.g., Scanlon's) to explain our results. We thank an anonymous reviewer for pointing this out.

Despite these limitations, this research provides strong experimental evidence in support of bargaining power influencing moral judgments, in line with recently proposed but currently understudied contractualist accounts of moral cognition.

References

Amasino, D. R., Pace, D. D., & van der Weele, J. (2023). Self-serving bias in redistribution choices: Accounting for beliefs and norms. *Journal of Economic Psychology*, *98*, Article 102654. https://doi.org/10.1016/j.joep.2023.102654

André, J., Fitouchi, L., Debove, S., & Baumard, N. (2023, May 10). *An evolutionary contractualist theory of morality*. PsyArXiv. https://doi.org/10.31234/osf.io/2hxgu

Aristotle. (2009). *The Nichomachean Ethics* (W. D. Ross & L. Brown, Eds.). Oxford University Press (Original work published c. 350 BC).

Babcock, L., & Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. *The Journal of Economic Perspectives*, *11*(1), 109–126. https://doi.org/10.1257/jep.11.1.109

Bacharach, S. B., & Lawler, E. J. (1981). Power and tactics in bargaining. *Industrial & Labor Relations Review*, *34*(2), 219–233. https://doi.org/10.1177/001979398103400204

Baumard, N. (2016). *The Origins of Fairness: How Evolution Explains our Moral Nature*. Oxford University Press.

Baumard, N., André, J.-B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences*, *36*(1), 59–78. https://doi.org/10.1017/S0140525X11002202

Binmore, K. (1994a). *Game Theory and the Social Contract—Volume I: Playing Fair*. MIT Press.

Binmore, K. (1994b). *Game Theory and the Social Contract—Volume II: Just Playing*. MIT Press.

Binmore, K. (2005). *Natural Justice*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195178111.001.0001

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of mice, men, and trolleys: Hypothetical judgment versus real-life behavior in trolley-style moral dilemmas. *Psychological Science*, *29*(7), 1084–1093. https://doi.org/10.1177/0956797617752640

Bruner, J., & O'Connor, C. (2016). Power, bargaining, and collaboration. In T. Boyer-Kassem, C. Mayo-Wilson, & M. Weisberg (Eds.), *Scientific Collaboration and Collective Knowledge: New Essays* (pp. 135–158). Oxford University Press.

Chater, N., Zeitoun, H., & Melkonyan, T. (2022). The paradox of social interaction: Shared intentionality, we-reasoning, and virtual bargaining. *Psychological Review*, *129*(3), 415–437. https://doi.org/10.1037/rev0000343

Crockett, M. J. (2013). Models of morality. *Trends in Cognitive Sciences*, *17*(8), 363–366. https://doi.org/10.1016/j.tics.2013.06.005

Cushman, F. A. (2013). Action, outcome, and value: A dual-system framework for morality. *Personality and Social Psychology Review*, *17*(3), 273–292. https://doi.org/10.1177/1088868313495594

Emerson, R. M. (1962). Power-dependence relations. *American Sociological Review*, *27*(1), 31–41. https://doi.org/10.2307/2089716

Everett, J. A., Pizarro, D. A., & Crockett, M. J. (2016). Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology: General*, *145*(6), 772–787. https://doi.org/10.1037/xge0000165

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

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T. (2012). What we say and what we do: The relationship between real and hypothetical moral choices. *Cognition*, *123*(3), 434–441. https://doi.org/10.1016/j.cognition.2012.02.001

Gauthier, D. (1986). *Morals by Agreement*. Oxford University Press.

Greene, J. D. (2014). *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*. Penguin.

Guala, F. (2013). Bargaining power and the evolution of un-fair, non-mutualistic moral norms. *Behavioral and Brain Sciences*, *36*(1), 92–93. https://doi.org/10.1017/S0140525X12000829

Hennig-Schmidt, H., Irlenbusch, B., Rilke, R. M., & Walkowitz, G. (2018). Asymmetric outside options in ultimatum bargaining: A systematic analysis. *International Journal of Game Theory*, *47*(1), 301–329. https://doi.org/10.1007/s00182-017-0588-4

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), Article 29. https://doi.org/10.1038/466029a

Hsu, L. C. (2008). Fairness and bargaining power in threshold public goods experiments. *Journal of Behavioral Decision Making*, *21*(2), 151–167. https://doi.org/10.1002/bdm.578

Kim, P. H., & Fragale, A. R. (2005). Choosing the path to bargaining power: An empirical comparison of BATNAs and contributions in negotiation. *Journal of Applied Psychology*, *90*(2), 373–381. https://doi.org/10.1037/0021-9010.90.2.373

Kuhn, J. W., Lewin, D., & McNulty, P. J. (1983). Neil W. Chamberlain: A retrospective analysis of his scholarly work and influence. *British Journal of Industrial Relations*, *21*(2), 143–160. https://doi.org/10.1111/j.1467-8543.1983.tb00127.x

Le Pargneux, A., Chater, N., & Zeitoun, H. (2024). Contractualist tendencies and reasoning in moral judgment and decision making. *Cognition*, *249*, Article 105838. https://doi.org/10.1016/j.cognition.2024.105838

Levine, S., Chater, N., Tenenbaum, J. B., & Cushman, F. A. (2024, June 20). *Resource-rational contractualism: A triple theory of moral cognition*. PsyArXiv. https://doi.org/10.31234/osf.io/p48t7

Levine, S., Kleiman-Weiner, M., Chater, N., Cushman, F. A., & Tenenbaum, J. B. (2024). When rules are over-ruled: Virtual bargaining as a contractualist method of moral judgment. *Cognition*, *250*, 105790. https://doi.org/10.1016/j.cognition.2024.105790

Levine, S., Kleiman-Weiner, M., Schulz, L., Tenenbaum, J. B., & Cushman, F. A. (2020). The logic of universalization guides moral judgment. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(42), 26158–26169. https://doi.org/10.1073/pnas.2014505117

Lois, G., & Riedl, A. (2022). Interplay between different forms of power and meritocratic considerations shapes fairness perceptions. *Scientific Reports*, *12*(1), Article 11428. https://doi.org/10.1038/s41598-022-15613-9

Malle, B. F. (2021). Moral judgments. *Annual Review of Psychology*, *72*(1), 293–318. https://doi.org/10.1146/annurev-psych-072220-104358

Mallucci, P., Wu, D. Y., & Cui, T. H. (2019). Social motives in bilateral bargaining games: How power changes perceptions of fairness. *Journal of Economic Behavior & Organization*, *166*, 138–152. https://doi.org/10.1016/j.jebo.2019.08.022

Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. (2014). Unwritten rules: Virtual bargaining underpins social interaction, culture, and society. *Trends in Cognitive Sciences*, *18*(10), 512–519. https://doi.org/10.1016/j.tics.2014.05.010

Nagel, T. (1995). *Equality and Partiality*. Oxford University Press. https://doi.org/10.1093/0195098390.001.0001

O'Connor, C. (2019). *The Origins of Unfairness: Social Categories and Cultural Evolution*. Oxford University Press. https://doi.org/10.1093/oso/9780198789970.001.0001

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press. https://doi.org/10.1017/CBO9780511807763

Ostrom, E. (2000). Collective action and the evolution of social norms. *The Journal of Economic Perspectives*, *14*(3), 137–158. https://doi.org/10.1257/jep.14.3.137

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *The American Political Science Review*, *86*(2), 404–417. https://doi.org/10.2307/1964229

Pinkley, R. L., Neale, M. A., & Bennett, R. J. (1994). The impact of alternatives to settlement in dyadic negotiation. *Organizational Behavior and Human Decision Processes*, *57*(1), 97–116. https://doi.org/10.1006/obhd.1994.1006

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, *118*(1), 57–75. https://doi.org/10.1037/a0021867

Rawls, J. (1971). *A Theory of Justice*. Harvard University Press. https://doi.org/10.4159/9780674042605

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org

Rode, J., & Le Menestrel, M. (2011). The influence of decision power on distributive fairness. *Journal of Economic Behavior & Organization*, *79*(3), 246–255. https://doi.org/10.1016/j.jebo.2011.02.003

Rustichini, A., & Villeval, M. C. (2014). Moral hypocrisy, power and social preferences. *Journal of Economic Behavior & Organization*, *107*, 10–24. https://doi.org/10.1016/j.jebo.2014.08.002

Starmans, C., Sheskin, M., & Bloom, P. (2017). Why people prefer unequal societies. *Nature Human Behaviour*, *1*(4), Article 0082. https://doi.org/10.1038/s41562-017-0082

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, *46*(1), 35–57. https://doi.org/10.1086/406755