SUPPLEMENT
Learning another's preference to punish increases one's own punitive behavior

Oriel FeldmanHall, A Ross Otto, Elizabeth A. Phelps

Experiment 1

Methods

Explanation of each option in the Justice Game.

According to the theory of retributive justice, an appropriate response to a fairness violation is to ensure that punishment is proportionate to the crime committed. Retributive justice can be dated back through recorded history, and is enshrined within most modern legal systems. These philosophies are formalized in psychological theory: if the punishment fits the crime, a person is deservingly punished proportionate to the moral wrong committed, often referred to as a 'just deserts' or deservingness principle (Carlsmith et al. 2002). In order to operationalize this in our task, we reasoned that reversing the Players' outcomes allows for the maximum punishment to be applied to Player A while also giving the maximum compensation to Player B. Moreover, reversing the Players' payouts results in Player A receiving what was initially assigned for Player B, and vice versa—a direct implementation of the 'just deserts' principle.

In lieu of punishing the criminal, justice can also be restored by providing monetary compensation to the victim (Weitekamp 1993). Research demonstrates that individuals can exhibit strong social preferences for equitable and efficient outcomes that increases the payouts of all recipients (Charness and Rabin 2002). Theories of fairness (Fehr and Schmidt 1999), and previous empirical work (FeldmanHall et al. 2014), argue that individuals have preferences that align with compensating the victim rather than punishing the perpetrator. We operationalized the 'Compensate' as increasing the victim's (Player B) monetary payout without decreasing Player A's payout (the Pareto efficient option). Finally, the option to Accept an offer from Player A reflects a classic option in the existing literature (Camerer 2003). When accepting an offer from Player A, Player B is typically agreeing to receive a smaller amount relative to what Player A apportions for him or herself.

Based on prior work employing multiple different variants of this task (FeldmanHall et al. 2014), we presented all three options (Accept, Reverse, Compensate) on every trial. This was done for the following reason: In one of the original versions (FeldmanHall et al. 2014), participants were randomly presented with only two options on any given trial, mirroring many classic experimental economics games that examine trade-offs between discrete choice pairs (Guth et al. 1982, Fehr and Gachter 2000). Player A—when making their offers—were not aware which two options would be available to Player B on any given trial, and thus could not pre-emptively decide to make strategic splits that may lead to maximizing the monetary pie since the option for Player B to Compensate was not available on every trial. In a second version (FeldmanHall et al. 2014), participants were presented with a plurality of options, such that all options were available on every trial. Despite these task differences, Player B's behavior was strikingly similar, revealing strong and systematic endorsement of the Compensate option (FeldmanHall et al. 2014).

Task Protocol.

To ensure task comprehension, participants had to correctly complete a quiz following the instructions. Once the quiz was correctly completed participants began the task by placing their hands on the keyboard on the following keys: S, D, F and a timer counted down from five before the task started. On each trial, the options 'compensate', 'accept', and 'reverse' (labeled here, but not presented to participants) were always displayed in a different order. After completing the entire task (all phases of the experiment), participants were probed on their strategies when the offer was relatively fair and when the offer was highly unfair when deciding for themselves. Specifically, participants were asked *"in your own words please describe your strategy for a scenario when Player A kept $.90 and offered $.10 to you"*. Responses included: *"I think it is best to get the highest possible money, even though Player A did not choose to pay out fairly"; "I thought he needs to see how it feels if someone does the same to him, then he might be more fair next time"; "I would be upset that they offered me an unfair bonus and would justify it by taking more"*.

Reinforcement Learning Model.

To investigate the process by which individuals learn about another's preferences, we leveraged a computational Reinforcement Learning framework. In total, we made three different predictions about how individuals might learn from others, resulting in three different Reinforcement-Learning (RL) models, each of which makes different assumptions about how feedback shapes the learning process. These models characterize how participants' responses are guided by estimates of the success of each response—that is, whether their response matches the preferences of the Receiver—which are continually updated by new feedback from the Receiver. The simplest model—termed the Basic Model—characterizes trial-by-trial learning without accounting for the fairness level of the offer at hand. Accordingly, on each trial t, a participant's action (Reverse, Compensate or Accept) was determined by the softmax rule:

$$P(reverse)_t = \frac{\exp(\beta * Q[reverse])}{exp\,[\beta * Q(reverse)] + exp\,[\beta * Q(compensate)] + \ exp\,[\beta * Q(accept)]}$$

where β is an inverse temperature (or sensitivity) parameter that governs the choice rules' sensitivity to value estimates (Q values; the estimated success of each action). Following each choice on trial t, the feedback received (1 if the participant's choice correctly mirrored the desired outcome of Receiver, and elsewise 0) was used to update Q values for the subsequent trial, Q_{t+1}. For example, if Reverse was the participant's response on trial t, the following update rule would apply:

$$Q(reverse)_{t+1} = Q(reverse)_t + \ \alpha[Q(reverse)_t - feedback_t]$$

where *feedback_t* denotes whether the participant's choice correctly reflected the feedback on the current trial t, and α is the learning rate. This update rule only updates the value of the chosen action, regardless of whether the action was correct. In this case, the assumption here is that a correct action is to Reverse if the participant is in the Punishment condition. The Q values for the two options that were not chosen (Compensate and Accept) are decayed as follows:

$$Q(compensate)_{t+1} = (1 - \alpha) * \ Q(compensate)_t$$
$$Q(accept)_{t+1} = (1 - \alpha) * \ Q(accept)_t$$

Effectively, the Basic Model maintains value estimates for the three possible responses, but ignores how unfairly the Receiver was treated, and is thus considered the simplest model. Following typical formulations of RL (Yechiam and Busemeyer 2005, Gershman et al. 2009, Otto et al. 2014), this model requires two free parameters: a learning rate (α), which governs the extent to which recent outcomes are weighted in updating estimates of action values, and a sensitivity parameter (β) which controls how sensitive participants' actions are to these values estimates.

In contrast, the Fairness Model assumes that people are sensitive to the magnitude of the fairness violation and might respond differently depending on whether the offer from Player A was somewhat fair [$6, $4] or highly unfair [$9, $1]. If it is the case that learning is sensitive to the extent of the fairness infraction, then such a model should do a better job of reflecting actual behavior then the Basic Model. To implement this sensitivity, action value estimates (Qs) are conditioned upon offer level, yielding 12 independent Q values (and like the Basic model, two free parameters). Accordingly, the probability of making, say, a Reverse response, is dependent upon the current offer type at trial t:

$$P(reverse \mid offer)_t$$
$$= \frac{\exp(\beta * Q[reverse|offer])}{exp\,[\beta * Q(reverse|offer)] + exp\,[\beta * Q(compensate|offer)] + \ exp\,[\beta * Q(accept|offer)]}$$

Additionally, updates to action values Q are conditioned on the offer type of the current trial. For example, if Reverse is chosen on trial t in response to offer type "*offer*", the following update would be applied:

$$Q(reverse \mid offer)_{t+1} = Q(reverse \mid offer)_t + \ \alpha[Q(reverse \mid offer)_t - feedback_t]$$

and the Q values for the other actions, contingent on that offer level, would be decayed as:

$$Q(compensate|offer)_{t+1} = (1 - \alpha) * \ Q(compensate|offer)_t$$

$$Q(accept|offer)_{t+1} = (1 - \alpha) * Q(accept|offer)_t$$

Finally, the Extended Fairness Model builds upon the Fairness Model, but with the addition of a separate learning rate for each offer type, yielding three separate learning rates, $\alpha_{reverse}$, $\alpha_{compensate}$, and α_{accept}, which are used to perform updates based on the offer type of the present trial. In this model, there are six free parameters that capture variable learning rates depending on how severe the fairness transgression is.

Our model-fitting procedure sought parameter values that maximized the log-likelihood of participants' choices given their previous choices and feedback. To compare goodness-of-fit across different models, we utilized the Bayesian information criterion (Schwarz 1978), defined as BIC = $-2 \times$ LL + k $\times$ log(n), where k is the number of free parameters in the model, *LL* is the log likelihood of the model given the participants' data, and n is the number of choices fit. Lower BIC values indicate better fit (summary statistics for best-fitting parameter values and goodness-of-fit measures are provided in the results section).

Results

Behavioral Results.

For each condition, we plot raw behavior (endorsement of Compensate or Reverse, depending on the condition) across all three phases of the task (Fig S1).
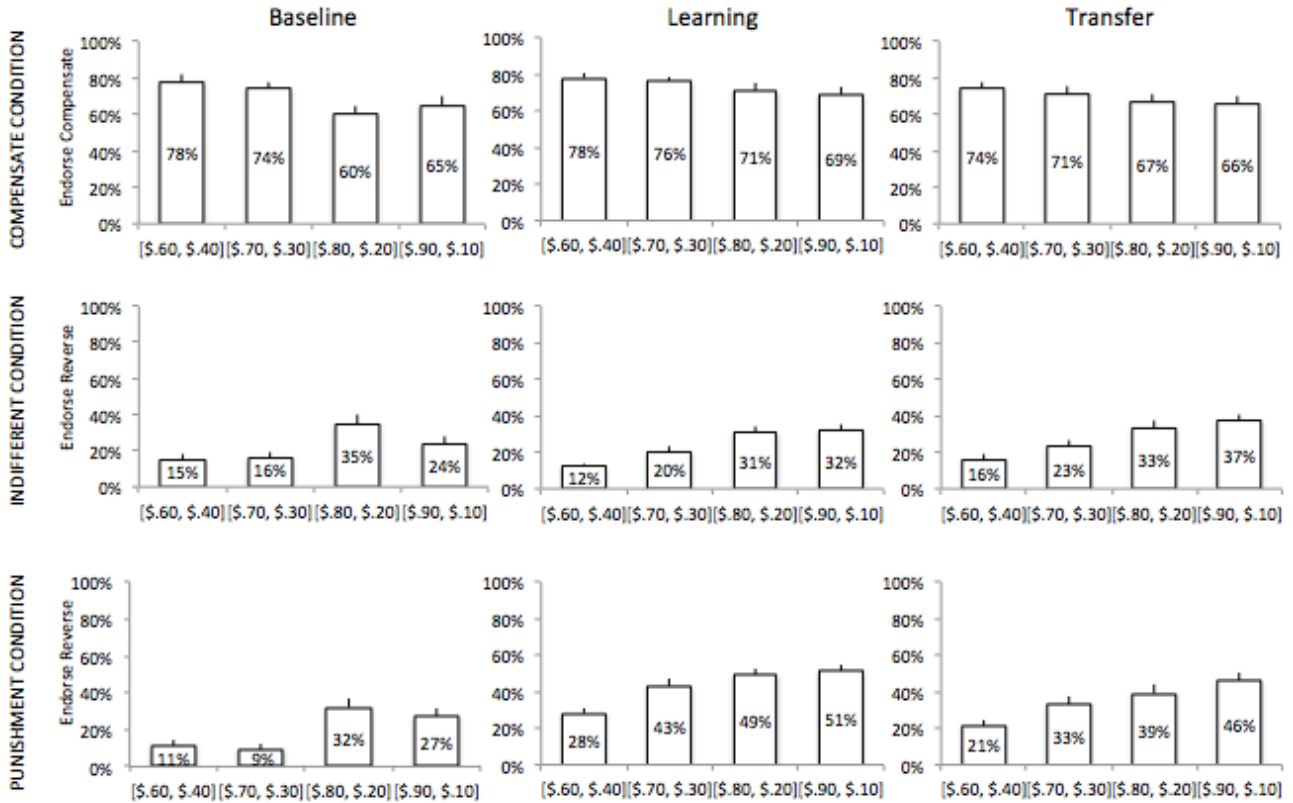
Fig S1 Experiment 1 | *Endorsement Rates broken down by each offer level for every phase (Baseline, Learning and Transfer) of the task. For the Compensate Condition, decisions to Compensate are plotted along the Y axis. For the Indifferent and Punishment Conditions, decisions to Reverse are plotted along the Y axis. Error bars reflect 1 SEM.*

Relationship between learning rate and transfer effects.

Estimated learning rates from the Fairness Model indicate participants also widely varied in how well they learned (Table S1). This raises the question of whether participants who exhibited more efficient learning (as operationalized by model goodness-of-fit (LL) since this parameter best captures learning) by consistently implementing the Receiver's feedback had the greatest transfer effects (greatest Response Δ). Results reveal that across all conditions, participant specific model fits predicted greater transmission of another's fairness preferences (all Ps<0.05, Table S2; NB: In the Indifferent condition there is a reverse correlation indicating no implementation).

In addition, in the Punishment condition—where we observed the greatest transfer of fairness preferences from the Receiver to the participant—participants who showed no behavioral

6

modifications in the Transfer phase (Response Δ≤0) had significantly worse model fits compared to participants who changed their behavior to match the Receiver's preferences (mean log likelihoods (LL) for participants with Δ≤0: -67.53 SD±26.11; mean LL for participants with Δ>0: -55.79 SD±25.91, independent t-test t(97)=-2.35, p=.02).

Table S1 | Model fits and parameter estimates for Each condition in Experiment 1.

Condition		Fairness Model Descriptives			
		Min	Max	Mean	SE
Compensate	Log Likelihoods	-91.12	-8.31	-39.46	3.02
	Alphas	.01	.74	.37	.01
Indifferent	Log Likelihoods	-90.62	-20.32	-54.90	2.45
	Alphas	.03	.42	.19	.007
Punishment	Log Likelihoods	-91.31	-8.37	-61.72	2.68
	Alphas	.01	.84	.32	.02

Table S2 | Spearman's correlations between Response Δ and Model Fits (Log likelihoods)

Conditions	Correlation Coefficient	P value
Compensate	.234	.02*
Indifferent	-.282	.005*
Punishment	.244	.015*

Relationship between magnitude of fairness violation and learning.

Does the capacity for learning depend on the magnitude of injustice (i.e. unfairness of the offer)? Research demonstrates that as fairness violations become more severe, acceptance of offers diminishes and the importance of rebalancing the scales of justice increases (Camerer 2003). In our Justice game, the magnitude of Player A's fairness offense was structured in stepwise increments, which allows us to explore two possible roles fairness violations play in shaping learning. One possibility is that learning is impervious to the degree of the infraction, and does not vary with respect to the magnitude of the offer. If this were the case, then similar transfer effects should be observed regardless of how unfairly one is treated.

A second possibility is that an individual's preferences might be sensitive to the fairness infraction. For example, fairness preferences may be more malleable and susceptible to being altered only if the infraction is opprobrious. This would be indicative that another's social preferences are most influential in guiding preferences for restoring justice after clearly flagrant fairness infractions (Camerer 2003). Alternatively, preferences might be most labile after a more morally ambiguous fairness infraction (Wiener et al. 1957, Hamm and Hoving 1969). If this were the case, it would indicate that fairness violations that grossly deviate from socially normative behavior elicit responses insensitive to social influence, while less offensive fairness infractions allow the individual to be more flexible in their learning responses for restoring justice.

We observed that the relationship between successful trial-and-error learning and subsequent matching of behavior to the punitive Receiver, was largely carried by ambiguous fairness violations (Spearman's correlation for [$.70, $.30] offer: LL x Response Δ: r=.30, p<0.001; Fig S2). This suggests that when the fairness violation is somewhat ambiguous, participants are better able to learn via trial-and-error about the Receiver's fairness preferences, resulting in greater acquisition and endorsement of punitive actions for themselves. In other words, how much another's preference for punishment was acquired, and how well the participant implemented or learned those preferences, were more tightly coupled after ambiguous fairness violations. This suggests that learning, and the resulting conformist moral behavior, requires some moral ambiguity in how unfairly one is treated. If the fairness violation is not egregious, people may have the perception of greater "wiggle room" to learn and implement another's preferences (Wiener 1957).
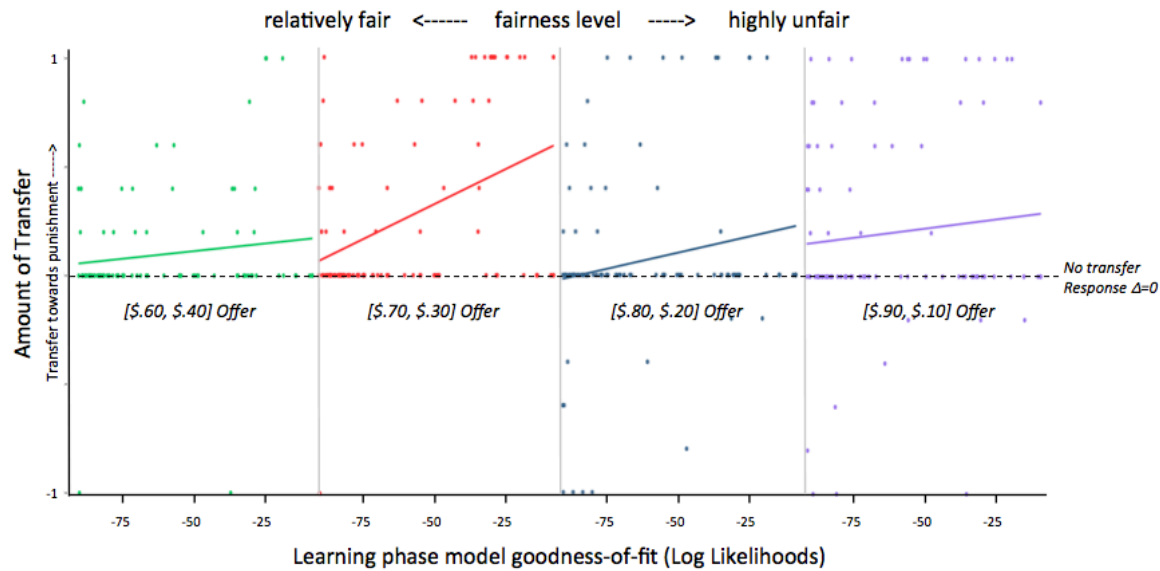
Fig S2 | Relationship between learning and transfer effects in Punishment Condition. *Across all conditions, overall model fits—i.e. successful learning—predict larger transmission of fairness preferences. Here we graph the degree to which participants increased their endorsement of Reverse for each offer level in the Punishment condition. The transfer effect is largely carried by individuals modifying their behavior most when the offers were somewhat morally ambiguous ($.70, $.30) and not too extreme.*

Experiment 2

Methods

Laboratory protocol.

In Experiment 2 it was important for participants to know that their decisions could impact the outcomes of others. Thus, participants were explicitly told a cover story in which they would be making decisions on behalf of past volunteers from the NYU community, and that future participants would also be making responses for them. This information was conveyed in the following way:

"Previously, we had people come into the lab and we had them act either as Player A or as Player B in this task. Their answers were recorded and have been input into this program. Their responses are what you are seeing today. Thus, you will be making decisions with real past participants, and your decisions will impact not only your own monetary payment, but their monetary payment as well. Payment will work as follows: At the end of the experiment, a trial will be randomly picked by the computer to be paid out. Based on that trial, all players will receive the money as you determined on that trial. We will pay you today and send a check to the

other participant depending on how you redistributed the money. Given that we use participants'
responses for future studies, we were also hoping you would allow us to use your responses to be
fed forward to the next participant. If so, we will collect those responses at the end of the task,
and also collect your name and address so we can email you a check. If you agree this will be
explained in more detail after the task."

At the end of the experiment, participants were also funneled debriefed about how well they believed the task. This debriefing is the typical way to discover whether the participants did not believe the manipulation, and therefore should not be included during analysis. They were told the following: *"As I explained in the beginning of the task, you played with multiple other people today. We have been running this study for a while, and while you were playing with real past participants, some people have wondered whether they were really playing with past volunteers. While you were, it is important for us to understand whether you ever doubted that. So, on a scale from 1 – 6 where 1 is had no doubts and 6 is filled with doubts, where would you put yourself?"* Dovetailing with the criteria used in past experiments (FeldmanHall et al. 2012, FeldmanHall et al. 2015, Murty et al. 2016), participants who report a 5 or 6 are typically not included in analysis for failing to believe the paradigm. Since the results of Experiment 2 hold with or without their inclusion, results are reported on the full sample (N=37).

Debriefing statements.

During debriefing, participants expressed their general strategies during the task. In some instances, participants explained that they tried to take the view of the Player B (who they were making a decision for) when making decisions for them. For example, in the following case, the participant states that they took the feedback from Player B when deciding on behalf of them, but did not do this when they themselves were the victim: *Sub1: "I found myself changing strategies. At first I just wanted to make everything equal. But then in C, it doesn't matter what I choose and A is being greedy, and the feedback influenced me that A is being greedy. So, after that I started reversing a lot more, unless he did 6-4. For that I would compensate. But 7-3 and up I liked to teach a lesson. When I was player B, I mostly compensated." Sub2: "If I felt player A was giving 9-1, I would punish Player A, and if it was 6-4, I would compensate. I was fairly*

similar if I was player B vs Player C. I only let player b's feedback influence me in 7-3 or 8-2 ratios when I was player C."

In other cases, however, participants explicitly reported that seeing what another did changed their own responses, resulting in a punitive transfer effect: *Sub3: "For the first part I started compensating mostly, and then after seeing some of the responses, it sort of changed some of my responses, and if player a was more fair I would compensate but if it was really unfair then I would reverse, and this kind of carried over when I was player B again."* However, in some of these instances, they were unaware that their decisions to punish more harshly in the Transfer phase may have been a result of learning about another's fairness preferences: *Sub 4: "I did the same for all- if it was 9 or 8 I would reverse, and if it were any of the others, I would compensate. I didn't really take player b's feedback. The second time I was player b, I reversed the 7-3 offers more than I did the first time I was player B, just because they annoyed me."*

Given that participants played with many different Player As, is also highly unlikely that the participants thought that responding more punitively would deter future bad behavior. Indeed, none of the debriefing statements suggests that the participants thought their responses might deter Player A from making unfair offers on the next round. For example, one subject said: *"When I was player B I generally just compensated myself. It didn't really make sense to do anything else for me. When I was player C, I generally reversed the outcome. I thought it strange a couple times that people accepted the outcome. When I was Player C, seeing if they wanted to reverse it, I guess it influenced my decision a little bit. I kind of understand wanting to punish other people. So I tried to do that for others, even if I didn't need to punish for myself."* Another said: *"I felt like I was in the shoes of Player B, and how I would feel if someone would propose an offer to me, and I don't think I would like it. Either way, I was not okay with Player A's choices, and I felt like this person was not a good person. So I found myself reversing a lot."* And another: *"I found myself changing strategies. At first I just wanted to make everything equal. But then in C, it doesn't matter what I choose and A is being greedy, and the feedback influenced me that A is being greedy. So, after that I started reversing a lot more, unless he did 6-4. For that I would compensate. But 7-3 and up I liked to teach a lesson. When I was player B, I mostly compensated."*

Results

Behavioral Results.

We plot raw behavior (endorsement of Reverse) across all three phases of the task (Fig S3).
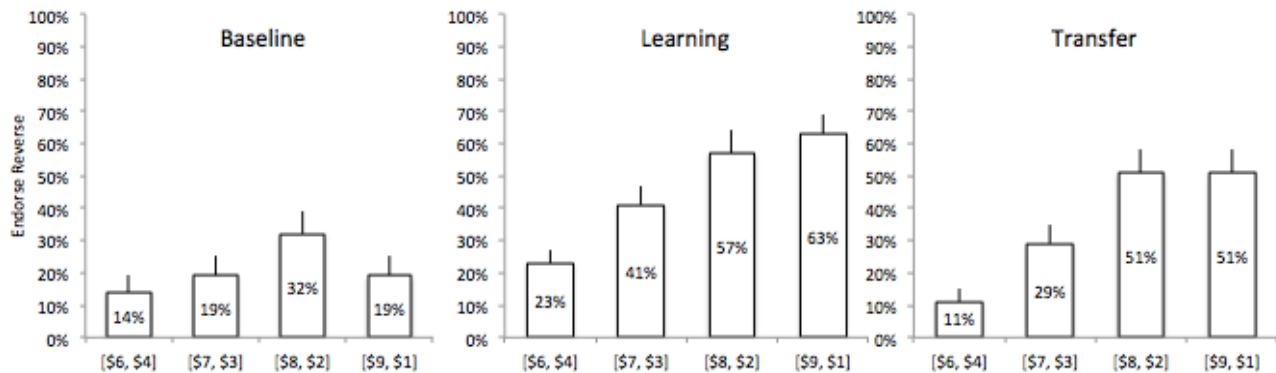

Fig S3 Experiment 2 *| Endorsement Rates of Reverse, broken down by each offer level for every phase of the task. Decisions to Reverse are plotted along the Y axis. Error bars reflect 1 SEM.*

Relationship between learning rate and transfer effects.

Comparing model fit, we again found evidence that the Fairness Model best characterizes overall learning (lowest BIC scores, Table S3). Estimated learning rates from the Fairness Model (Table S4) revealed that those with no behavioral modifications in the Transfer phase (Response Δ≤0) had worse model fits compared to participants who changed their behavior to match the Receiver's preferences (mean LL for participants with Δ≤0: -69.52 SD±22.93; mean LL for participants with Δ>0: -58.63 SD±18.25, independent t-test, t(35)=-1.60, p=.11). Mirroring the findings from the first experiment, participant specific model fits—i.e. successful learning— predicted greater transmission of fairness preferences (Spearman's correlation LL x Response Δ: r=.49, p=.002, Table S4), and this was strongest for ambiguous fairness violations (Spearman's correlation for [$7, $3] offer: LL x Response Δ: r=.39, p=.019, Fig S4). Together, these findings further support the idea that successful trial-and-error learning results in a greater shift towards endorsing punishment as a means of restoring justice.

Table S3| Summary of Model Goodness-of-fit Metrics

RL Models	Mean (SE) BIC scores by Condition
	Punishment
Basic Model	*161.01 (9.3)*
*Fairness Model**	*136.61 (8.1)*
Extended Fairness Model	*146.85 (8.3)*

Table S4 | Experiment 2 (Punishment Condition) Model fits and parameter estimates for Fairness Model

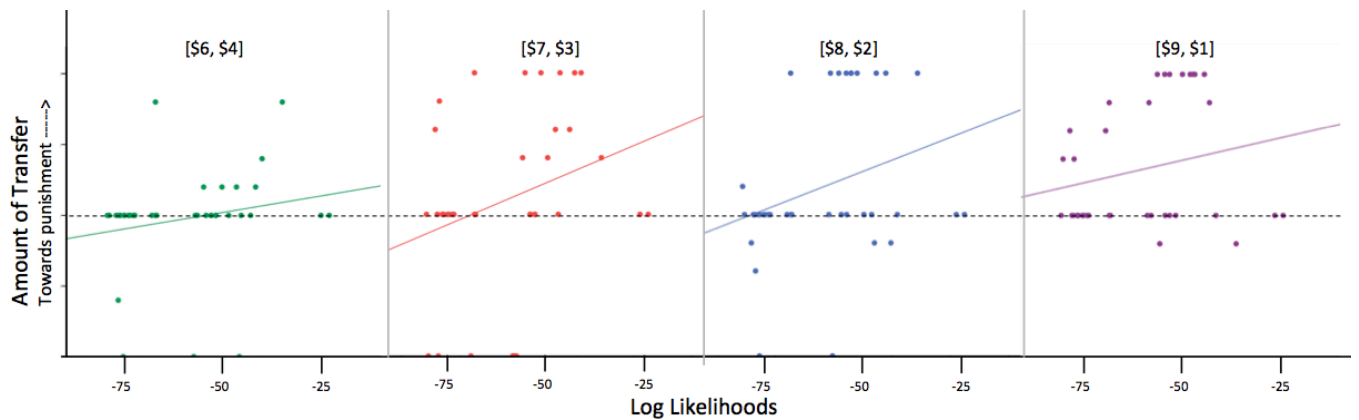| Condition		Fairness Model Descriptive				Correlation between Response Δ & Model Fits (Log likelihoods)
		Min	*Max*	*Mean*	*SE*	
Punishment	*Log Likelihoods*	*-91.33*	*-15.14*	*-63.93*	*3.47*	*r=0.485* *p=0.002**
	Alphas	*.01*	*.89*	*.36*	*.033*	

Fig S4 | Relationship between learning and transfer effects in Punishment Condition in Experiment 2*. In the Punishment condition, overall model fits—i.e. successful learning—predict larger transmission of fairness preferences. Here we graph the degree to which participants increased their endorsement of Reverse for each offer level. The transfer effect is largely carried by individuals modifying their behavior most when the offers were somewhat morally ambiguous.*

Experiment 3

Results

Relationship between learning rate and transfer effects.

A comparison of the reinforcement learning models revealed the Fairness Model again outperformed the other models (mean Fairness Model BIC=80.64 SE=3.84; mean Extended Fairness Model BIC=92.02 SE=1.97), indicating that participants were sensitive to the extent of fairness violation, but did not exhibit different learning rates for each offer type.

To explore the relationship between learning rates and how much people changed their punitive preferences, we tested whether participants without behavioral modifications in the Transfer phase (Response Δ≤0) had worse model fits compared to participants who changed their behavior to match the Receiver's preferences in the Transfer phase (Response Δ>0). Results reveal that those who did not change their preferences had worse model fits (mean LL for participants with Δ≤0: -45.93 SD±10.79; mean LL for participants with Δ>0: -34.08 SD±13.5, independent t-test, t(49)=2.4, p=.024). Corroborating this result, successful learning (indexed by participant-specific model goodness-of-fit metrics) predicted greater transmission of fairness preferences (Spearman's correlation LL x Response Δ: r=.36, p=.013, Table S5), and this was significant across all types of fairness violations (Spearman's correlation, all Ps<.01). Together, these findings further support the generality of the phenomenon that successful trial-and-error learning and implementation of another's punitive preferences results in a greater shift towards endorsing punishment as a means of restoring justice.

Table S5 | Model fits and parameter estimates for Fairness Model Experiment 3 (UG)

Condition		*Fairness Model Descriptives*				Correlation between Response Δ & Model Fits (Log likelihoods)
		Min	*Max*	*Mean*	*SE*	
Punishment	*Log Likelihoods*	*-59.39*	*-5.97*	*-35.94*	*1.92*	*r=0.36* *p=0.013**
	Alphas	*0*	*.79*	*.32*	*.027*	

Experiment 4

Results

Behavioral Results.

For both the active and passive learning conditions, we plot raw behavior (endorsement of Reverse) across all three phases of the task (Fig S5).
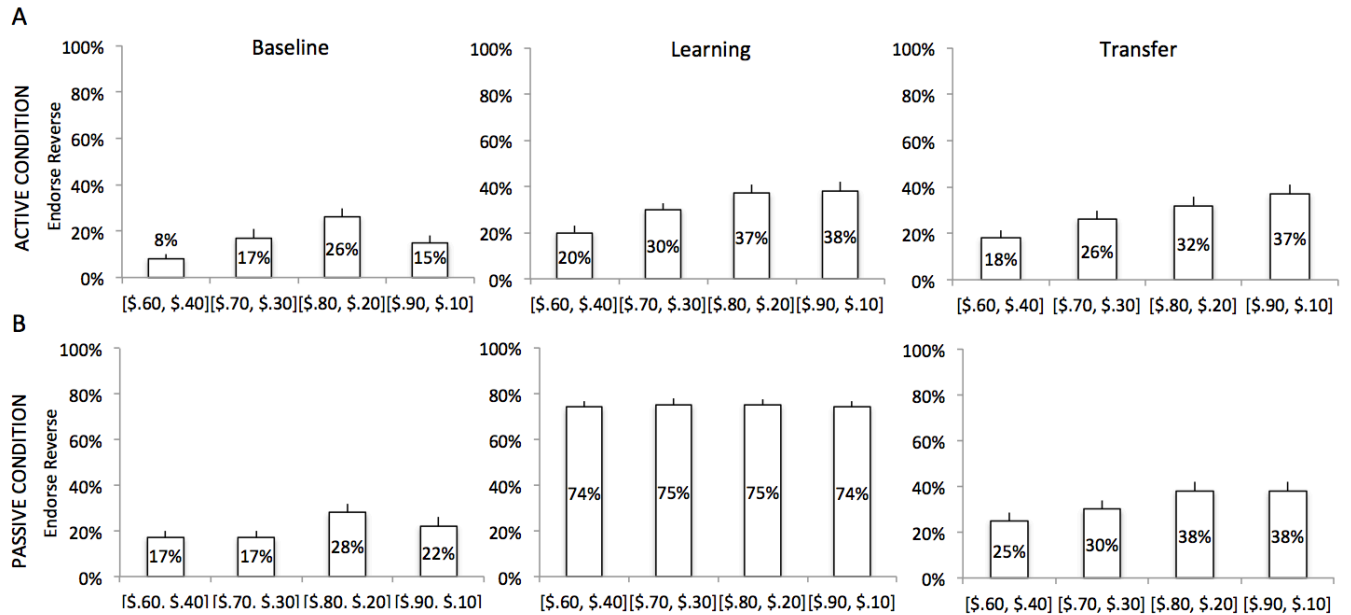

Fig S5 Experiment 4 | A) Active Condition. *Endorsement Rates of Reverse, broken down by each offer level for every phase of the task.* ***B) Passive Condition.*** *Endorsement Rates of Reverse, broken down by each offer level for every phase of the task. Note that during the Learning Phase, decisions to Reverse reflect that subjects were asked to report what the Receiver had just selected, which can be considered an attention check. Decisions to Reverse are plotted along the Y-axis. Error bars reflect 1 SEM.*

Relationship between learning rate and transfer effects.

As we had done in the other experiments, we examined whether, during the active learning condition, acquiring the preferences of another requires more successful learning and implementation. Comparing model fit, we again found evidence that the Fairness Model best characterizes overall learning in the active learning condition (lowest BIC scores, Table S6). For the active learning condition, estimated learning rates from the Fairness Model revealed that those with no behavioral modifications in the Transfer phase (Response $\Delta \leq 0$) had worse model fits compared to participants who changed their behavior to match the Receiver's preferences (mean LL for participants with $\Delta \leq 0$: -72.4 SD±18.6; mean LL for participants with $\Delta > 0$: -62.7 SD±26.8, independent t-test, t(93)=-2.10, p=.04). When we examined the relationship between Response Δ and Model Fits, we found that those who acquired the punitive preferences of the Receiver in the Transfer phase had lower LLs (correlation between LL x Response Δ: r=.31,

15

p<0.001), and as in the previous experiments, this effect was carried by the [$.70, $.30] offer (LL x Response Δ: r=.32, p=0.002).

Table S6 | Summary of Model Goodness-of-fit Metrics in Experiment 4

RL Models	Mean (SE) BIC scores *Active Learning*	Mean (SE) BIC scores *Passive Learning*
Basic Model	*174.01 (5.2)*	*106.01 (4.7)*
Fairness Model*	*145.38 (4.6)*	*113.58 (3.8)*
Extended Fairness Model	*156.85 (4.7)*	*125.36 (3.9)*

Table S7 | Model fits and parameter estimates for Fairness Model in Experiment 4

Condition		**Fairness Model Descriptive**			
		Min	*Max*	*Mean*	*SE*
Active Learning	*Log Likelihoods*	*-91.11*	*-9.54*	*-68.31*	*2.33*
	Alphas	*.01*	*.70*	*.27*	*.02*

For the passive learning condition, we found no differences between those with no behavioral modifications and those who acquired another's preferences (mean LL for participants with Δ≤0: -50.2 SD±18.5; mean LL for participants with Δ>0: -55.2 SD±20.3, independent t-test, t(98)=-1.3, p=.20). It is important to note, however, that in the passive learning task model fits reflect how well a participant correctly responded (i.e., accurately) with what the Receiver actually decided to do on that trial (accuracy scores for each offer level hovered around 80%). We also observed no relationship between Response Δ and Model Fits (correlation between LL x Response Δ: r=-.05, p=.59).

REFERENCES:
Camerer, C. (2003). Behavioral game theory : experiments in strategic interaction. New York, N.Y.
Princeton, N.J., Russell Sage Foundation ;
Princeton University Press.

Carlsmith, K. M., J. M. Darley and P. H. Robinson (2002). "Why do we punish? Deterrence and just deserts as motives for punishment." Journal of Personality and Social Psychology **83**(2): 284-299.

Charness, G. and M. Rabin (2002). "Understanding social preferences with simple tests." Quarterly Journal of Economics **117**(3): 817-869.

Fehr, E. and S. Gachter (2000). "Cooperation and punishment in public goods experiments." American Economic Review **90**(4): 980-994.

Fehr, E. and K. M. Schmidt (1999). "A theory of fairness, competition, and cooperation." Quarterly Journal of Economics **114**(3): 817-868.

FeldmanHall, O., D. Mobbs, D. Evans, L. Hiscox, L. Navrady and T. Dalgleish (2012). "What we say and what we do: The relationship between real and hypothetical moral choices." Cognition **123**(3): 434-441.

FeldmanHall, O., C. M. Raio, J. T. Kubota, M. G. Seiler and E. A. Phelps (2015). "The Effects of Social Context and Acute Stress on Decision Making Under Uncertainty." Psychol Sci **26**(12): 1918-1926.

FeldmanHall, O., P. Sokol-Hessner, J. J. Van Bavel and E. A. Phelps (2014). "Fairness violations elicit greater punishment on behalf of another than for oneself." Nature Communications **5**.

Gershman, S. J., B. Pesaran and N. D. Daw (2009). "Human reinforcement learning subdivides structured action spaces by learning effector-specific values." J Neurosci **29**(43): 13524-13531.

Guth, W., R. Schmittberger and B. Schwarze (1982). "An Experimental-Analysis of Ultimatum Bargaining." Journal of Economic Behavior & Organization **3**(4): 367-388.

Hamm, N. H. and K. L. Hoving (1969). "Conformity of Children in an Ambiguous Perceptual Situation." Child Development **40**(3): 773-&.

Murty, V. P., O. FeldmanHall, L. E. Hunter, E. A. Phelps and L. Davachi (2016). "Episodic memories predict adaptive value-based decision-making." Journal of Experimental Psychology-General **145**(5): 548-558.

Otto, A. R., W. B. Knox, A. B. Markman and B. C. Love (2014). "Physiological and behavioral signatures of reflective exploratory choice." Cognitive Affective & Behavioral Neuroscience **14**(4): 1167-1183.

Schwarz, G. (1978). "Estimating the Dimension of a Model." The Annals of Statistics **6**(2): 461-464.

Weitekamp, E. (1993). "Reparative Justice." European Journal of Criminal Policy and Research **1**(1): 70-93.

Wiener, M., J. T. Carpenter and B. Carpenter (1957). "Some Determinants of Conformity Behavior." Journal of Social Psychology **45**(2): 289-297.

Wiener, M., Carpenter, J. T. & Carpenter, B. (1957). "**Some Determinants of Conformity Behavior.**" Journal of Social Psychology **45**: **289-297**

Yechiam, E. and J. R. Busemeyer (2005). "Comparison of basic assumptions embedded in learning models for experience-based decision making." Psychon Bull Rev **12**(3): 387-402.