

Preprint – Author’s Original Manuscript

This manuscript represents the version submitted for publication and has not been peer-reviewed.

A peer-reviewed version of this work has since been published in *Basic and Applied Social Psychology*.¹

Reciprocity and Generalized Reciprocity: How Unfairness Affects Allocations in First and Second Interactions with Wrongdoers and Uninvolved Parties

Joanna Rudzińska-Wojciechowska*¹, Jerzy Wojciechowski², Katarzyna Sekścińska²

¹Centre for Economic Psychology and Decision Sciences, Kozminski University,

Jagiellońska 57, 03-301, Warsaw, Poland

²Faculty of Psychology, University of Warsaw, Stawki 5/7, 00-183, Warsaw, Poland

*Corresponding author

E-mail address: jrudzinska@kozminski.edu.pl

Acknowledgements

Compliance with Ethical Standards

All procedures were conducted in accordance with the ethical standards of the institutional and/or research committees and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The Ethics Board of Kozminski University approved the studies.

¹ Note on publication status

A peer-reviewed version of this manuscript has been published as:
Rudzinska-Wojciechowska, J., Wojciechowski, J., & Sekścińska, K. (2025).
Reciprocity and Generalized Reciprocity: How Unfairness Affects Allocations in First and Second Interactions with Wrongdoers and Uninvolved Parties.
Basic and Applied Social Psychology.
<https://doi.org/10.1080/01973533.2025.2572630>

Conflict of Interest

The authors declare that there are no relevant financial or non-financial competing interests to report.

Open Data

Complete data can be found at:

<https://repod.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/EC7T9I>

Funding: This work was supported by the National Science Centre, Poland, grant number: UMO-2020/39/D/HS4/01880. The funding source had no involvement in the studies' designs; the collection, analysis, and interpretation of data; the writing of this report; or in the decision to submit this article for publication.

AI usage statement

Model used: OpenAI ChatGPT (GPT-4)

Extent of usage: The AI tool was used to support language editing and to rephrase selected sentences for clarity and naturalness. It was also used to check the overall consistency and structure of the manuscript during revision. No content (theory, hypotheses, results, or interpretations) was generated by AI.

Abstract:

Most research on unfairness responses examines the first decision after a norm violation. We studied behavior across two interactions in a two-stage game (N = 680). Participants received an unfair offer, then allocated money to either the same person (the wrongdoer) or a new, uninvolved person. Later, they made a second allocation to the wrongdoer or another stranger. Participants gave less to wrongdoers than to new partners, especially if they had not previously had the opportunity to retaliate. When they had such an opportunity, this effect diminished. These findings suggest that direct retaliation may reduce punitive motivation, whereas generalized responses do not. The study clarifies how fairness concerns and reciprocity unfold across interactions and highlights the limits of displaced responses.

Keywords: reciprocity, direct reciprocity, generalized reciprocity, unfairness, inequality

Common wisdom, supported by research findings, suggests that people tend to respond to unfair treatment by reciprocating or by passing unkindness to others. For example, if person A treats person B unfairly, person B is likely to treat person A unfairly in return (Falk & Fischbacher, 2006). However, when direct retaliation is not possible, person B may instead direct their negative behavior toward an uninvolved third party, person C. This last scenario illustrates how unfairness can propagate through social interactions, a phenomenon well-documented in numerous studies (Gray et al., 2014; Liu et al., 2015; Stanca, 2009; Wu et al., 2015). Despite extensive research on the spread of unfairness beyond the original interaction most studies focus on single interactions within this chain. Prior studies have examined how individuals react to a single instance of unfairness, but fewer have explored behavioral patterns across multiple interactions. Specifically, we examine whether retaliating against the wrongdoer reduces further negative reciprocity or if it persists in subsequent encounters.

Furthermore, how does prior interaction with an uninvolved third party influence later treatment of the original wrongdoer? These questions are important for understanding how responses to unfairness evolve over time, particularly in contexts where repeated social exchanges shape cooperation and trust.

1. Literature Overview:

People strongly dislike inequity and are willing to punish those who treat them unfairly, even at a cost (Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999; Henrich et al., 2006; Johansson & Svedsäter, 2009; Norton & Ariely, 2011). This tendency is widely observed in social interactions, where individuals respond negatively to perceived injustice. Individuals who experience unfair treatment often respond with unkindness toward the person responsible for the wrongdoing (Falk & Fischbacher, 2006). This type of situation is known as direct reciprocity (Nowak & Sigmund, 2005). However, in many cases, direct retaliation is not possible. When individuals direct negative behavior toward an uninvolved third party after experiencing unfair treatment, this is referred to as generalized negative reciprocity (Pfeiffer et al., 2005). This mechanism can lead to a broader diffusion of unfairness in social networks, creating a 'chain of unfairness' where negative treatment extends beyond the original interaction (Gray et al., 2014; Liu et al., 2015; Stanca, 2009; Wu et al., 2015). Understanding how these forms of reciprocity unfold over multiple interactions is critical for assessing how unfair treatment shapes long-term social behavior.

One of the most common methods for studying reciprocity is through economic games, particularly the dictator game and the ultimatum game. The ultimatum game requires one participant to propose a division of a sum of money, which the other participant can either accept or reject, with both receiving nothing in the case of rejection (Güth et al., 1982; see Guala, 2008 for a review). In contrast, the dictator game removes the recipient's power to reject: one player determines the distribution of money or other goods that a second player

must accept (Kahneman et al., 1986; see Guala & Mittone, 2010, for a review and Engel, 2011, for a meta-analysis). These games have been widely used to investigate reactions to fairness and unfairness, allowing researchers to analyze both direct and generalized reciprocity in controlled settings. By applying these paradigms, previous studies have demonstrated how unfair treatment influences subsequent decisions, offering a framework for understanding the mechanisms underlying negative reciprocity.

Direct reciprocity is often studied using the ultimatum game. The outcomes of this game offer valuable insights not only into the concept of direct reciprocity but also into the notion of fairness. While standard game theory suggests that participants should propose minimal offers (Guala, 2008), empirical results reveal that fairness is a deeply ingrained behavioral norm. By default, people tend to split outcomes equally, often in a 50%/50% fashion (Messick & Schell, 1992; Thaler & Camerer, 1995). Moreover, even though it is rational to always accept the offered resources, a majority of responders typically reject offers below 20%–30% of the total, accepting personal loss to punish unfairness (Camerer & Thaler, 1995).

While direct reciprocity involves retaliation against the initial wrongdoer, generalized reciprocity reflects how unfair treatment can shape future interactions with uninvolved third parties. Generalized reciprocity is most frequently explored through two-stage games where individuals who receive an offer in the first stage subsequently become dictators in the second stage and interact with a new, uninvolved partner. For instance, Gray et al. (2014) manipulated fairness by assigning participants equal, greedy, or generous divisions of money or labor from anonymous partners. They then examined how participants distributed new resources to another anonymous recipient. Participants who received greedy treatment tended to pass on this behavior, allocating less money and more unpleasant labor to the next person compared to those who were treated fairly. Similar findings were obtained by Wu et al.

(2015), where participants repeatedly received fair or unfair offers in the ultimatum game and later had the chance to retaliate in dictator games with unrelated players. Again, individuals exposed to unfair divisions allocated fewer resources to new partners. These findings suggest that individuals do not merely react to direct unfairness but also propagate unfair treatment within social networks, reinforcing cycles of negative reciprocity.

Both direct and generalized reciprocity involve retaliatory or harmful behaviors in response to perceived unfair treatment, which can negatively impact relationships. While these forms of reciprocity may seem similar, there are key differences between them. The main distinction lies in their social scope: direct reciprocity occurs between two parties—the wrongdoer and the retaliator—making it a personal exchange. In contrast, generalized reciprocity affects a wider group, often involving individuals who were not involved in the initial harm. Additionally, while there is extensive research on the mechanisms behind direct negative reciprocity, studies on generalized reciprocity remain relatively scarce.

Direct reciprocity is typically driven by a sense of fairness, where individuals retaliate against those who have wronged them (Fehr & Gächter, 2000). This behavior is seen as a way to punish defectors and avoid an inferior social status (Sigmund et al., 2002). From an evolutionary perspective, direct reciprocity is considered adaptive and prosocial because individuals sacrifice resources to punish unfair behavior, which helps maintain cooperation in social contexts such as friendships and markets (Fehr & Henrich, 2003; Hoffman et al., 1998). Moreover, several studies highlight the role of emotions in negative reciprocity, showing that negative emotions such as anger can motivate retaliatory decisions (Bosman & Van Winden, 2002; Sanfey et al., 2003). Other research suggests that people derive satisfaction from punishing defectors, sometimes prioritizing emotional fulfillment over monetary gain (Ben-Shakhar et al., 2007).

In the case of generalized reciprocity, motives rooted in personal retaliation may not fully explain this phenomenon. Hence, much of the research on the mechanisms of generalized reciprocity has focused on the role of emotions. For example, Gray et al. (2014), investigated whether reducing negative emotions could lessen the tendency for generalized reciprocity. In their study, participants who received unfair treatment were more generous to the next participants after engaging in a positive affect intervention—rating cartoons—compared to the controls. Similarly, Strang et al. (2016) found that writing letters to the wrongdoer was an effective way to regulate emotions, leading to more generous behavior toward a third party.

On the other hand, some studies suggest that the mechanisms behind generalized reciprocity may not necessarily be tied to emotions or emotion regulation. For example, in a study by Schnedler & Stephan (2020), participants who were treated unfairly and then asked to write letters to their wrongdoers gave more assets to others in a subsequent dictator game compared to participants who were simply asked to wait for three minutes. Interestingly, however, the act of writing a letter did not improve the self-reported happiness of participants who had been treated unkindly. Moreover, even those who had been treated kindly also gave more to third parties when given the opportunity to write letters. The authors propose that writing letters may activate prosocial thinking regardless of whether the individual experienced kindness or unfairness beforehand. In another study that explored whether emotional mechanisms alone can explain generalized reciprocity (Schnedler, 2022), participants who were treated unfairly reported stronger negative emotions than those treated kindly, yet their generosity toward unrelated third parties did not significantly differ from that of the more kindly treated group.

In summary, although both forms of reciprocity seem to have much in common at first glance, relatively few studies have examined the mechanisms behind sharing decisions in the

context of generalized negative reciprocity. These studies suggest that this form of reciprocity may not be fully explained by emotional reactions alone, highlighting the need for further investigation to better understand the underlying processes. One way to extend our knowledge of how direct and generalized negative reciprocity differ is to compare their intensity in response to the same unfair offer in a controlled setting.

Although the literature consistently demonstrates the effects of both direct and generalized negative reciprocity, the specific comparison between their intensities—particularly in response to identical unfair stimuli—has not been sufficiently explored. Existing studies bring mixed results: some suggest that individuals respond with similar intensity to both forms (Herne et al., 2013), while others suggest that the relationship between received and sent amounts is weaker in cases of generalized reciprocity (Ben-Ner et al., 2004). However, these studies often do not clearly distinguish whether the observed differences in reciprocity intensity stem from the directness of retaliation or from other contextual factors. Notably, these studies included responses to fair, generous, and unfair offers in their analyses, despite growing evidence that positive and negative reciprocity are distinct mechanisms (e.g., Yamagishi et al., (2012), suggesting that these should be analyzed separately to avoid conflating different motivational processes.

To our knowledge, no studies have directly compared the intensity of direct and generalized reciprocity in response to the same unfair treatment. This gap is significant not only from a methodological perspective but also for understanding the fundamental psychological and social mechanisms that drive reciprocity. By systematically comparing these two forms of reciprocity in an identical context of unfairness, we can examine whether differences in their intensity stem from the directness of retaliation or from distinct underlying processes governing fairness-related decision-making. This comparison helps clarify whether personal retaliation (direct reciprocity) is more intense than broader, less personalized

responses (generalized reciprocity), thereby deepening our understanding of how unfairness spreads in social interactions. Accordingly, we pose the following research question:

RQ1: After experiencing the same unfair treatment, are there significant differences in subsequent allocations to the wrongdoer and to an uninvolved third party?

Drawing on the literature reviewed above, we predict that participants are more likely to treat the wrongdoer unkindly, deviating from norms of fairness, when compared to an uninvolved third party. One key difference between direct and generalized reciprocity is that only direct reciprocity provides an opportunity to punish the person responsible for the unfair treatment. Punishment is a fundamental mechanism for enforcing fairness norms and discouraging unfair behavior in repeated interactions (Fehr & Gächter, 2000). Since generalized reciprocity does not allow for direct punishment of the wrongdoer, individuals may be less motivated to act unkindly toward an uninvolved third party.

H1: Allocations to the wrongdoer will be lower than those to an uninvolved third party.

Furthermore, much of the existing research has focused on participants' initial reactions after experiencing unfair treatment, leaving open the question of how behavior evolves in subsequent interactions. Specifically, if a wronged individual has an opportunity to reciprocate against the wrongdoer, will they continue to treat this person unfairly in the second encounter? Conversely, if they have a chance to treat a third person badly in the first interaction after being treated unfairly, how will they respond when they later face the original wrongdoer? Although prior studies have examined how unfairness propagates in social interactions (Gray et al., 2014; Wu et al., 2015), it remains unclear whether a single opportunity to reciprocate or pass on unfairness to a new individual is sufficient to prevent further negative behavior in subsequent interactions.

In this study, we investigate participants' behavior in a second interaction following an initial experience of unfairness. After receiving an unjust allocation, participants will either be given the chance to reciprocate against the wrongdoer or transfer the unfairness to an uninvolved third party in their first interaction. We then examine their behavior in a second interaction, where they are paired either with the same wrongdoer or with another uninvolved third party. This allows us to test whether initial interactions shape future allocation decisions, specifically whether an early opportunity to retaliate reduces the likelihood of further negative reciprocity. This raises two research questions:

RQ2: Among individuals who had the chance to pay unfairness forward in the first interaction, are there significant differences in allocations in the second interaction when comparing those who interact with the wrongdoer and those who interact with a third, uninvolved person?

RQ3: Among individuals who had the opportunity to reciprocate in the first interaction, are there significant differences in allocations in the second interaction when comparing those who interact with the wrongdoer and those who interact with a third, uninvolved person?

Despite the lack of direct studies on this topic, previous research on reciprocity and fairness norms provides theoretical grounds for predicting how individuals may behave in subsequent interactions. In both cases—whether individuals have the opportunity to treat a third party unfairly or directly retaliate against the wrongdoer—there are some similarities. Both situations allow individuals to respond to the initial unfair treatment by reducing allocations to another person. However, only direct reciprocity provides an opportunity to punish the actual wrongdoer, which may serve not only as retaliation but also as a means of enforcing fairness norms (Fehr & Gächter, 2002).

One possible explanation for how these behaviors unfold comes from catharsis theory, which suggests that retaliating against a transgressor can reduce the motivation for further punitive actions (Bushman, 2002; Denzler, Förster, & Liberman, 2009). This idea is based on the assumption that expressing negative reactions—whether through direct retaliation or displacement onto a third party—provides psychological relief, reducing the need for further punitive behavior (Geen & Quanty, 1977). However, prior research indicates that not all forms of retaliation are equally effective in reducing subsequent punitive tendencies. Studies have shown that venting aggression on an unrelated target does not necessarily diminish future punitive impulses and may, in some cases, even increase retaliatory motivation (Bushman et al., 1999).

If a similar pattern applies to responses to unfairness, individuals who initially allocate fewer resources to an uninvolved third party may not perceive this as sufficient retribution, leaving them inclined to treat the wrongdoer more harshly when encountering them later. Conversely, those who had the opportunity to directly reciprocate against the original wrongdoer may feel that they have adequately responded to the initial unfair treatment, reducing their motivation to continue punishing in later interactions.

A complementary explanation concerns the role of punishment in enforcing social norms. Revenge is not solely an emotional reaction but also a mechanism for maintaining cooperation and signaling that unfair behavior is unacceptable (Gollwitzer & Denzler, 2009; Miller, 2001; Vidmar, 2005). Prior research suggests that when individuals have an opportunity to punish a transgressor, they may perceive this as sufficient enforcement of fairness norms, reducing the need for further punishment in future interactions (Gollwitzer & Denzler, 2009; Fehr & Gächter, 2002). Conversely, if they do not get this chance, the perceived injustice remains unresolved, which may lead to continued punitive behavior when they later face the original wrongdoer (Denzler et al., 2009; Vidmar, 2005).

Based on this, we expect that if participants do not have the opportunity to retaliate directly in the first interaction, they will allocate fewer resources to the wrongdoer in the second interaction compared to an uninvolved third party. Conversely, when participants do have an opportunity to directly reciprocate against the wrongdoer in the first interaction, they may perceive their response as sufficient, making subsequent allocations more comparable between the wrongdoer and a third party. Therefore, we hypothesize only that:

H2: Among individuals who did not have the opportunity to reciprocate in the first interaction following the unfair treatment, allocations to the wrongdoer in the second interaction will be lower than those to an uninvolved third party.

2. The current study

2.1 Study overview:

The study aimed to examine how individuals allocate resources in two consecutive interactions after experiencing unfair treatment, specifically comparing their behavior toward the original wrongdoer and an uninvolved third party. In the experiment, all participants initially received unfair treatment in a dictator game. Since our focus was on the effects of prior unfairness, we did not include a control group without such treatment. Instead, we compared behaviors across conditions that differed in opportunities for direct or generalized reciprocity. Subsequently, it was the participant's turn to play the role of dictator. In this round, some of them had the opportunity to reciprocate to the wrongdoer, while others could only direct their response toward an unrelated third person. We then compared these allocations to examine whether an initial opportunity for direct retaliation influenced subsequent decisions. Finally, all participants engaged in a second interaction, either with another new person or with the individual who had treated them unfairly in the initial interaction. Our goal was to observe participants' decisions in this interaction to determine

whether there were differences in allocations between these conditions. These experimental groups were designed to directly test our hypotheses regarding the intensity of direct versus generalized reciprocity (H1) and the effects of initial interaction type on subsequent behavior (H2), as outlined in the previous section. Informed consent was obtained from all participants. The Ethics Board of Kozminski University approved the study.

2.2 Methods:

Participants: 680 participants took part in the study (365 females, 314 males, 1 other) aged 18 - 65 ($M = 42.36$; $SD = 13.57$). The study was conducted using an online Polish participant panel called Ariadna. Participants received a fixed number of points for participating (20 points) and additional points based on their performance in experimental tasks (ranging from 0 to 23). These points could be redeemed on the panel's website for a reward of their choice.

Materials and Procedure: In this study, participants engaged in three rounds of the dictator game, each involving 100 units of Experimental Dollars to be divided between the Dictator and the Recipient. At the beginning of the study, participants were informed that they would make three financial decisions in pairs. They also learned that before each decision, a computer would randomly pair participants, and one person in the pair would be chosen randomly to distribute 100 Experimental Dollars between themselves and their partner without consulting them. However, to maintain maximum control over experimental variables and allow for an experimental procedure, it was necessary to pre-determine the roles in each round. Consequently, in the first round, all participants acted as Recipients and received the same unequal treatment, with an unequal share of Experimental Dollars (split of 30%/70%). In the subsequent two rounds, all participants became Dictators and were asked to share Experimental Dollars either with the person who acted as a Dictator in the first round (referred to as the "wrongdoer") or with a new person, depending on experimental conditions (see Figure 1). To clarify:

- Participants in the first group (Wrongdoer - Wrongdoer, WW, $n = 177$) played with the wrongdoer in both the second and third rounds.
- Participants in the second group (Wrongdoer - New, WN, $n = 173$) played with the wrongdoer in the second round and with a new person in the third round.
- Participants in the third group (New – Wrongdoer, NW, $n = 161$) played with a new person in the second round and with the wrongdoer in the third round.
- Participants in the fourth group (New – New, NN, $n = 169$) played with a new person in the second round and another new person in the third round.

The number of Experimental Dollars assigned to the Receivers in the second and third rounds served as the dependent variable.

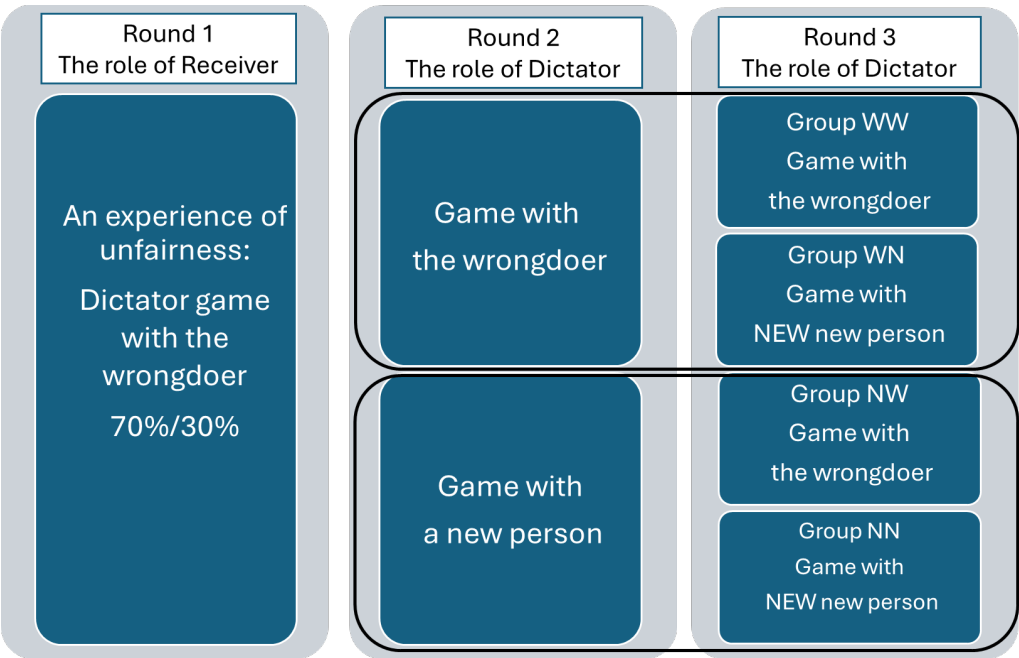


Figure 1: Experimental groups in the study

Throughout all stages of the game, the history of decisions was displayed, allowing participants to infer whether they were playing with the wrongdoer or a new person based on

prior allocations. The study was incentivized, with the Experimental Dollars accumulated by participants in their account after the three decisions being converted into additional points (10 Experimental Dollars = 1 point), which were added to the points awarded for study participation. At the end of the study, the participants were fully debriefed.

2.3 Results

Comparison of the allocations to the wrongdoer and a new person in the first interaction after an experience of unfairness (Round 2).

Descriptive statistics of the allocations made in the first interaction after an experience of unfairness made by participants who played the dictator game with the wrongdoer and those who played with a new person are presented in Table 1, while Figure 2 depicts distributions of the Experimental Dollars in the two groups.

Table 1. Descriptive Statistics: allocations in the first interaction after an experience of unfairness

	allocations	
	to the wrongdoer	to a new person
N	350	330
Median	40	49.5
Mean	37.143	39.779
Std. Deviation	17.131	16.088
Minimum	0	0
Maximum	100	100

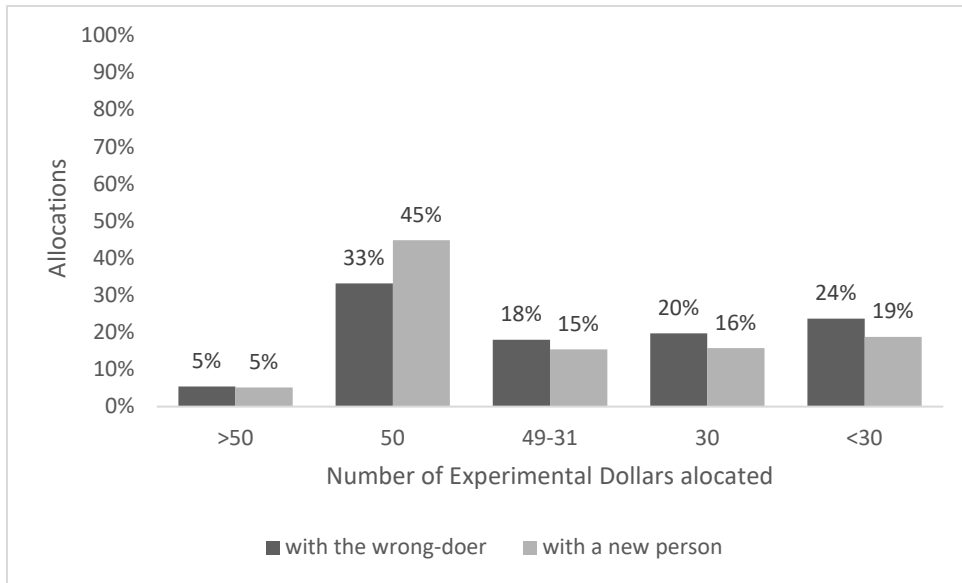


Figure 2. The distribution of Experimental dollars allocated in a game with the wrongdoer and a new partner in the first interaction after an experience of unfairness.

A Shapiro-Wilk test was performed considering a strong concentration of allocations of 50 Experimental Dollars. It showed that the distribution of allocations to the partner departed significantly from normality ($W = 0.9, p < 0.01$). Based on this outcome, a Mann-Whitney test was performed to evaluate whether allocations differed by the experimental group. The results indicated that people who played with the wrongdoer allocated significantly fewer Experimental Dollars to their partners than those who played with a new person, $U = 51304, p = 0.009, r_{rb} = -0.112$, see Figure 3.

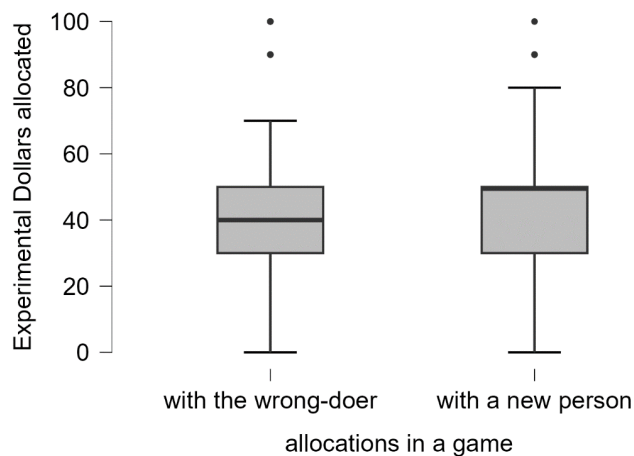


Figure 3. A comparison of experimental groups in the distribution of Experimental Dollars a game with the wrongdoer and a new partner in the first interaction after an experience of unfairness.

Seeking further insights into allocation decisions made by participants from the two groups, we examined whether the groups differed in the frequency of allocations below an equal split (below 50 Experimental Dollars) and equal to or lower than 30 Experimental Dollars (equal or lower than the original offer). Allocations that fall below 50% deviate from a 50/50 split and could be seen as unfair. Therefore, this analysis shows how often unfair allocations occur in both groups. The frequency of allocations at or below 30% represents the number of people in both groups who chose to either respond with the same treatment they received or worse, or to pay forward the same or worse treatment. The results of χ^2 analysis indicate that allocations below 50 Experimental Dollars were more frequent in the game with the wrongdoer (61.43%) than with a new person (50%), $\chi^2(1) = 8.998$, $p = 0.003$. Similar results were obtained when it comes to allocations equal to 30 Experimental Dollars or below; they were more frequent in a group that played with the wrongdoer (43.43%) than in a group that played with a new person (34.545%), $\chi^2(1) = 5.628$, $p = 0.018$.

Comparison of the allocations to the wrongdoer and the new person in the second interaction after an experience of unfairness (Round 3).

Further analyses compare allocation decisions in interactions with either the wrongdoer or a new person, focusing separately on two groups: those who played with the wrongdoer in Round 2 (WW vs. WN) and those who played with a new person (NW vs. WN). It's important to note that other comparisons (e.g., WN vs. NN) could not be made because there is more than one difference between the experimental conditions. Specifically, these conditions differ both in the identity of the interaction partner (wrongdoer vs. stranger) and in the type of interaction that occurred in the second round (direct vs. generalized reciprocity). Such

confounds would make it difficult to attribute differences in allocations to a specific factor.

Therefore, we focused only on comparisons that isolate a single experimental manipulation.

Allocations to the wrongdoer and the new person in the interaction preceded by interaction with the wrong-doer.

Descriptive statistics of the allocations made in the second interaction after an experience of unfairness, preceded by the opportunity to reciprocate, made by participants who played the dictator game with the wrongdoer and those who played with a new person, are presented in Table 2, while Figure 4 depicts distributions of the Experimental Dollars in the groups.

Table 2. Descriptive Statistics allocations in the second interaction after an experience of unfairness preceded by interaction with the wrong-doer.

	allocations in the second interaction	
	to the wrongdoer (WW)	to a new person (WN)
N	177	173
Median	40	40
Mean	37.011	38.162
Std. Deviation	18.387	17.972
Minimum	0	0
Maximum	100	100

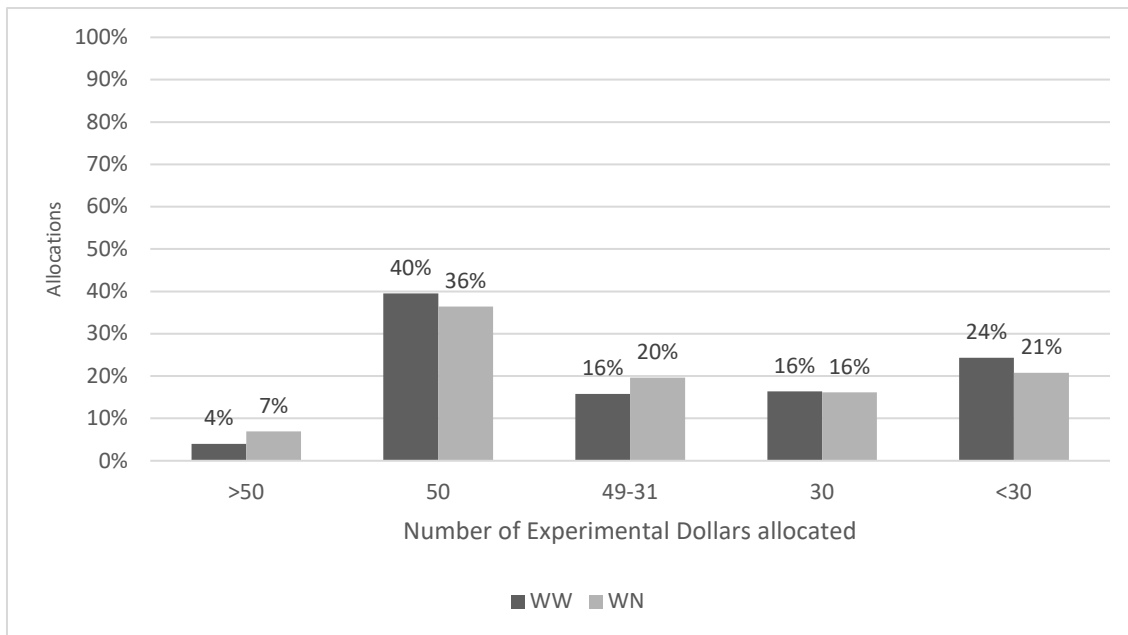


Figure 4. The distribution of Experimental dollars allocated in a game with the wrongdoer and a new partner in the first interaction after an experience of unfairness proceeded by interaction with the wrong-doer.

A Mann-Whitney test was performed to evaluate whether allocations differed between WW and WN groups. The results showed no differences in the number of experimental Dollars allocated to the new person and to the wrongdoer among people who played with a wrongdoer in the first interaction after an experience of unfairness $U = 14822,5$; $p = 0.595$, $r_{rb} = -0.032$, see Figure 5.

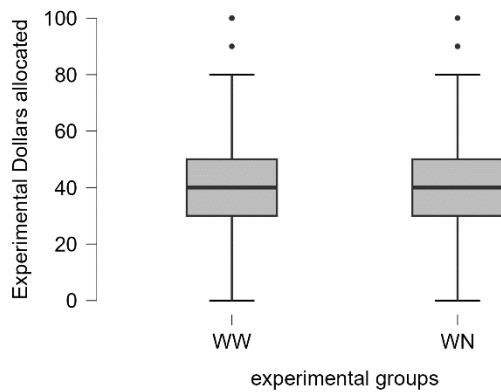


Figure 5. A comparison of experimental groups that played with a wrong doer in the first interaction after unfair treatment in the distribution of Experimental Dollars in a game with the wrongdoer (WW) and a new partner (WN).

Again, χ^2 analyses were conducted to analyze the differences in the frequency of allocations that are below an equal split. The frequency of allocations below 50 Experimental Dollars, and equal to or lower than 30 Experimental Dollars were compared. The results of χ^2 analysis indicate that among people who played with a wrong-doer in the first interaction after an experience of unfairness there is no significant difference in allocations below 50 Experimental Dollars between a group that played with the wrongdoer (50.51%) and a group that played with a new person (49.49%), $\chi^2(1) = 0.0008$, $p = 0.977$. Similar result were obtained when it comes to allocations equal to 30 Experimental Dollars or below (wrongdoer: 52.94 %, a new person: 47,06%, $\chi^2(1) = 5$, $p = 0.480$.)

Allocations to the wrongdoer and the new person in the interaction proceeded by interaction with the new person.

Descriptive statistics of the allocations made in the second interaction after an experience of unfairness, preceded by the opportunity to pay unfairness forward, made by participants who played the dictator game with the wrongdoer and those who played with a new person, are

presented in Table 3, while Figure 6 depicts distributions of the Experimental Dollars in the groups.

Table 3. Descriptive Statistics allocations in the second interaction after an experience of unfairness proceeded by interaction with the new person.

	allocations in the second interaction	
	to the wrongdoer (NW)	to a new person (NN)
N	161	169
Median	40	50
Mean	36.776	40.444
Std. Deviation	18.821	18.162
Minimum	0	0
Maximum	100	100

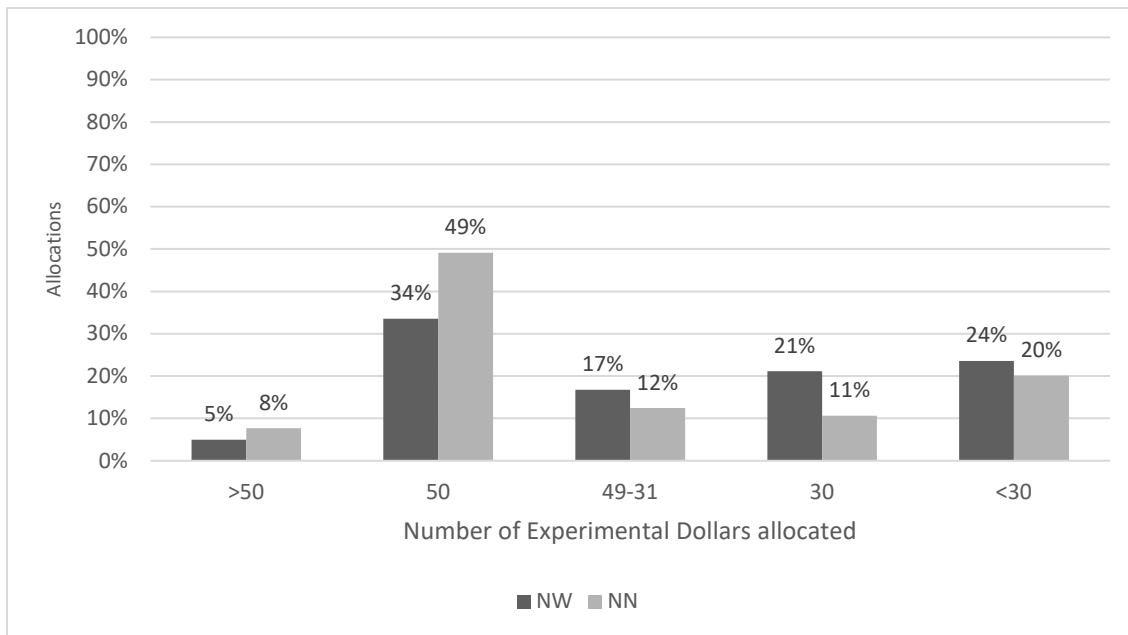


Figure 6. The distribution of Experimental dollars allocated in a game with the wrongdoer and a new partner in the second interaction after an experience of unfairness proceeded by interaction with the new person.

A Mann-Whitney test was performed to evaluate whether allocations differed between NN and NW groups. The results showed a significant difference in the number of experimental Dollars allocated to the new person and to the wrongdoer among people who played with a new person in the first interaction after an experience of unfairness $U = 11335,5$; $p = 0,006$, $r_{rb} = - 0.167$, see Figure 6.

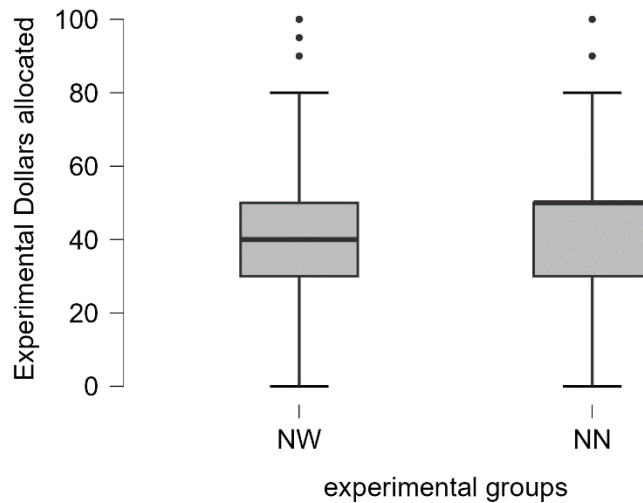


Figure 7. A comparison of experimental groups that played with a new person in the first interaction after unfair treatment in the distribution of Experimental Dollars in a game with the wrongdoer (NW) and a new partner (NN).

Again, χ^2 tests were performed to examine differences in the frequency of allocations below an equal split. Specifically, the frequency of allocations under 50 Experimental Dollars and those equal to or less than 30 Experimental Dollars were compared.

The χ^2 analysis results indicate that among people who played with a new person in the first interaction after an experience of unfairness, allocations below 50 Experimental Dollars were more frequent in a group that played with the wrongdoer (57.56%) than in a group that played with a new person (42.44%), $\chi^2(1) = 11.06$, $p < 0.001$. Similar results were obtained for allocations of 30 Experimental Dollars or less, with higher frequencies in interactions with the wrongdoer (58.07%) compared to a new person (41.94%), $\chi^2(1) = 6.842$, $p = 0.009$.

3. Discussion

The Study aimed to observe and compare allocation decisions made in the first and second interactions following experiences of unfair treatment, in the encounters with the wrongdoer and with a third, unrelated person.

The first research question (RQ1) examined whether participants would allocate resources differently to the wrongdoer versus an uninvolved third party after receiving identical unfair treatment. The results indicate that participants who engaged in direct reciprocity (i.e., allocated to the original wrongdoer) gave significantly fewer resources than those who engaged in generalized reciprocity (i.e., allocated to a new, uninvolved person). In this group, offers below an equal split and equal to or lower than the offers received from the wrongdoer are more frequent. These results align with Ben-Ner et al.'s (2004) who also observed a weaker correlation between the amounts received and sent in generalized reciprocity compared to direct reciprocity. However, they contrast with Herne's et al. (2013) findings, which reported no differences between these types of reactions. Notably, both studies measured the strength of reciprocal reactions in response to offers that varied not only in terms of unfairness but also fairness or generosity. A key distinction in the present study is that we examined responses to identical unfair offers, providing more precise insights into the differences between direct and generalized reciprocity following unfair treatment. The results help distinguish between personalized and broader social exchange dynamics.

Research questions 2 and 3 (RQ2 and RQ3) addressed whether prior interaction with the wrongdoer or a third party in the first post-unfairness decision would influence subsequent allocations in the second round. By focusing on second interactions, the study goes beyond the immediate consequences of unfairness and sheds light on the dynamics of retaliation and fairness in subsequent social encounters—an area largely unexplored in prior research. We restricted comparisons in the second interaction to groups that differed in only one critical

factor—whether participants had previously interacted with the wrongdoer or a stranger—to ensure that any observed differences could be clearly interpreted.

The results showed no significant differences in how participants allocated resources between the wrongdoer and a new person in cases where the wrongdoer was encountered in the first interaction. At this stage, there was no difference in generalized and direct reciprocity, neither in the median allocations nor in the frequency of offers below an equal split or equal to or lower than the initial unfair offer. This result aligns with previous findings (Denzler et al., 2009; Gollwitzer & Denzler, 2009) suggesting that revenge must be directed at the wrongdoer to reduce future aggressive impulses. Participants who directly retaliated against the wrongdoer in the first interaction did not treat the wrongdoer worse than a new person in the second interaction, indicating that the desire for revenge had been fulfilled through the initial act. While this is consistent with catharsis theory, it is important to note that emotional states were not directly measured, and such interpretations remain speculative.

In contrast, significant differences were found in the group that first interacted with a new, uninvolved person. In this group, participants allocated less to the wrongdoer than to the uninvolved person during the second interaction. The median allocation for direct reciprocity was lower than for generalized reciprocity and offers below an equal split or equal to or lower than the initial unfair offer were more frequent when participants played with the wrongdoer than when they interacted with the new person. Consistent with previous studies (Bushman, Baumeister, & Stack, 1999), our results show redirecting negative responses toward an unrelated third party does not diminish the motivation to later punish the original offender. Participants who lacked the opportunity to directly retaliate against the wrongdoer continued to treat them harshly in the second encounter. This finding also aligns with research indicating that when the desire for revenge is not addressed in a direct manner, the need for retaliation may persist and re-emerge in future interactions (e.g., Denzler et al., 2009).

An interesting observation from the distribution charts and frequency tables is that despite facing an unequal split, half of the participants who interacted with a new person allocated half or more of the endowment to them. This fraction increased to over 57% in the second interaction with a new person, where participants opted for an equal or generous split. Even in interactions with the wrongdoer, a significant proportion (38%) transferred half or more of the endowment, regardless of whether the encounter occurred in the first or second interaction. This finding raises questions about individual differences in fairness sensitivity and suggests a need for future research on resilience in social decision-making. One possible explanation for the relatively high proportion of equal splits is that the study used Experimental Dollars rather than real currency. However, a meta-analysis of dictator games shows that participants tend to allocate more when handling real money than hypothetical amounts (Engel, 2011). This suggests that the observed levels of fairness may, if anything, be underestimated—though this interpretation remains speculative and requires further testing.

3.1 Limitations:

While the results of this study provide valuable insights, several limitations should be acknowledged. First, the stakes in the game were relatively low. Although research has shown that people are sensitive to unfairness and willing to punish at a cost, regardless of the size of the stakes (Cameron, 1999; Munier & Zaharia, 2002) and despite that even hypothetical rewards yield similar results to real ones (Locey et al., 2011), a replication of the study using real, earned (Danková & Servátka, 2015) and substantial rewards would allow for greater confidence in the obtained results. Additionally, the study only examines one specific instance of unfairness (a 30/70 split). Although this ratio is considered unfair, there could be even more unjust scenarios, such as 0/100. Future studies should test whether the observed effects hold in situations of greater inequality. Finally, while the study focused specifically on reactions in response to unequal split, further studies could include additional conditions, such as fair and

generous treatments, to broaden our knowledge of reciprocity and generalized reciprocity in consecutive interactions. Another limitation is that the study was not preregistered. Although the experimental design and hypotheses were determined in advance based on prior literature, preregistration would improve the transparency and credibility of future replications.

3.2 Contributions

This study offers several contributions to the literature on reciprocity and social behavior. First, it offers a novel comparison between direct and generalized reciprocity in response to identical unfair stimuli, providing new insights into how individuals react differently in personal versus broader social contexts. Second, by examining second interactions, the study moves beyond the immediate consequences of unfairness to explore how behaviors evolve over time, an area largely underexplored in prior research. Third, the results suggest that some individuals demonstrate psychological resilience in the face of unfair treatment, opening new avenues for research on the factors that buffer against the spread of negative behavior.

Conclusion

In conclusion, the findings from this study provide insights into the nuanced dynamics of reciprocity following unfair treatment, particularly in second interactions where participants had the opportunity to either retaliate against the wrongdoer or treat a third party unfairly. Our findings show that people react more strongly when they can respond directly to unfairness, compared to when they deal with someone new. Additionally, the findings suggest that direct retaliation effectively reduces the desire for further revenge, whereas venting on an unrelated third party is insufficient to satisfy this need. Future research should continue exploring these dynamics, particularly in more extreme conditions of unfairness and with larger, real-world stakes.

Constraints on Generality

This study investigated allocation behavior following unfair treatment in economic games, focusing on differences between direct and generalized reciprocity across two consecutive interactions. Although the findings were robust within our experimental design, several constraints on generality should be considered.

Participants. The sample consisted of adult volunteers recruited via a commercial online research panel in Poland. While this offers greater diversity than student samples, it still reflects a WEIRD population. Cultural and socioeconomic factors may affect the generalizability of the findings, particularly in how individuals perceive and respond to fairness violations or norm violations. The results may not generalize to older or more culturally diverse populations without further replication.

Materials and procedures. The study employed a modified dictator game with a fixed unfair offer (30/70 split) and interactions involving either the original wrongdoer or a new, uninvolved person. Although this paradigm effectively captures core dynamics of fairness and reciprocity, the use of hypothetical currency (Experimental Dollars) instead of real money may have affected participants' engagement with the task. Nevertheless, prior meta-analyses suggest that differences between hypothetical and incentivized dictator games are limited, and hypothetical settings may even underestimate generous behavior (Engel, 2011).

Context. The experiment was conducted online and anonymously, simulating interactions in digital or deindividuated environments. While this reflects a common decision-making context in the modern world, it may limit generalizability to situations where social cues, face-to-face interaction, or reputational concerns are more salient. The absence of real-time

feedback or communication may also influence how participants process unfair treatment and formulate their responses.

Robustness across methods. We expect that the observed differences between direct and generalized reciprocity are likely to generalize across small variations in procedure, such as the size of the stake or timing between interactions. However, more substantial changes to the experimental context—such as introducing identifiable recipients, varying the level of unfairness, or including group membership—may significantly alter the pattern of responses. These factors could influence both the salience of norm violations and the motivation to retaliate or act fairly.

Conclusion. The current findings likely generalize to anonymous, economically framed dyadic interactions following mild norm violations. However, broader generalizations to more complex social contexts and populations require further empirical testing.

Bibliography:

- Ben-Ner, A., Putterman, L., Kong, F., & Magan, D. (2004). Reciprocity in a two-part dictator game. *Journal of Economic Behavior and Organization*, *53*(3), 333–352.
<https://doi.org/10.1016/j.jebo.2002.12.001>
- Ben-Shakhar, G., Bornstein, G., Hopfensitz, A., & van Winden, F. (2007). Reciprocity and emotions in bargaining using physiological and self-report measures. *Journal of Economic Psychology*, *28*(3), 314–323. <https://doi.org/10.1016/j.joep.2007.02.005>
- Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, *90*(1), 166–193. <https://doi.org/10.1257/aer.90.1.166>
- Bosman, R., & Van Winden, F. (2002). Emotional hazard in a power-to-take experiment. *Economic Journal*, *112*(476), 147–169. <https://doi.org/10.1111/1468-0297.0j677>
- Bresin, K., & Gordon, K. H. (2013). Aggression as affect regulation: Extending catharsis theory to evaluate aggression and experiential anger in the laboratory and daily life. *Journal of Social and Clinical Psychology*, *32*(4).
<https://doi.org/10.1521/jscp.2013.32.4.400>
- Bushman, B. J. (2002). *Does Venting Anger Feed or Extinguish the Flame? Catharsis, Rumination, Distraction, Anger, and Aggressive Responding*. *28*(6), 724–731.
<https://doi.org/10.1177/0146167202289002>
- Bushman, B. J., Baumeister, R. F., & Stack, A. D. (1999). Catharsis, aggression, and persuasive influence: Self-fulfilling or self-defeating prophecies? *Journal of Personality and Social Psychology*, *76*(3). <https://doi.org/10.1037/0022-3514.76.3.367>

- Camerer, C., & Thaler, R. H. (1995). Anomalies: Ultimatums, Dictators and Manners. *Journal of Economic Perspectives*, 9(2), 209–219. <https://doi.org/10.1257/jep.9.2.209>
- Cameron, L. A. (1999). Raising the stakes in the ultimatum game: Experimental evidence from Indonesia. *Economic Inquiry*, 37(1), 47–59. <https://doi.org/10.1111/j.1465-7295.1999.tb01415.x>
- Danková, K., & Servátka, M. (2015). The house money effect and negative reciprocity. *Journal of Economic Psychology*, 48. <https://doi.org/10.1016/j.joep.2015.02.007>
- Denzler, M., Förster, J., & Liberman, N. (2009). How goal-fulfillment decreases aggression. *Journal of Experimental Social Psychology*, 45(1), 90–100. <https://doi.org/10.1016/J.JESP.2008.08.021>
- Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, 14(4), 583–610. <https://doi.org/10.1007/s10683-011-9283-7>
- Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54(2), 293–315. <https://doi.org/10.1016/J.GEB.2005.03.001>
- Fehr, E., & Gächter, S. (2000). Fairness and Retaliation: The Economics of Reciprocity. *Ssrn*, 14(3), 159–181. <https://doi.org/10.2139/ssrn.229149>
- Fehr, E., & Henrich, J. (2003). Is Strong Reciprocity a Maladaptation? In *Genetic and Cultural Evolution of Cooperation*.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3). <https://doi.org/10.1162/003355399556151>
- Geen, R. G., & Quanty, M. B. (1977). The Catharsis of Aggression: An Evaluation of a Hypothesis. *Advances in Experimental Social Psychology*, 10(C). [https://doi.org/10.1016/S0065-2601\(08\)60353-6](https://doi.org/10.1016/S0065-2601(08)60353-6)

- Gollwitzer, M., & Denzler, M. (2009). What makes revenge sweet: Seeing the offender suffer or delivering a message? *Journal of Experimental Social Psychology*, 45(4), 840–844.
<https://doi.org/10.1016/J.JESP.2009.03.001>
- Gray, K., Ward, A. F., & Norton, M. I. (2014). Paying it forward: Generalized reciprocity and the limits of generosity. *Journal of Experimental Psychology: General*, 143(1), 247–254.
<https://doi.org/10.1037/a0031047>
- Guala, F. (2008). Paradigmatic Experiments: The Ultimatum Game from Testing to Measurement Device. *Philosophy of Science*, 75(5), 658–669.
<https://doi.org/10.1086/594512>
- Guala, F., & Mittone, L. (2010). Paradigmatic experiments: The Dictator Game. *Journal of Socio-Economics*, 39(5), 578–584. <https://doi.org/10.1016/j.socec.2009.05.007>
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3(4).
[https://doi.org/10.1016/0167-2681\(82\)90011-7](https://doi.org/10.1016/0167-2681(82)90011-7)
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardaroas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesoronol, C., Marlowe, F., Tracer, D., & Ziker, J. (2006). Costly punishment across human societies. *Science*, 312(5781), 1767–1770.
<https://doi.org/10.1126/science.1127333>
- Herne, K., Lappalainen, O., & Kestilä-Kekkonen, E. (2013). Experimental comparison of direct, general, and indirect reciprocity. *Journal of Socio-Economics*, 45.
<https://doi.org/10.1016/j.socec.2013.04.003>

- Hoffman, E., McCabe, K. A., & Smith, V. L. (1998). Behavioral foundations of reciprocity: Experimental economics and evolutionary psychology. *Economic Inquiry*, *36*(3).
<https://doi.org/10.1111/j.1465-7295.1998.tb01719.x>
- Johansson, L. O., & Svedsäter, H. (2009). Piece of cake? Allocating rewards to third parties when fairness is costly. *Organizational Behavior and Human Decision Processes*, *109*(2), 107–119. <https://doi.org/10.1016/j.obhdp.2009.02.001>
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the Assumptions of Economics. *The Journal of Business*, *59*(S4). <https://doi.org/10.1086/296367>
- Liu, P. P., Safin, V., Yang, B., & Luhmann, C. C. (2015). Direct and indirect influence of altruistic behavior in a social network. *PLoS ONE*, *10*(10).
<https://doi.org/10.1371/journal.pone.0140357>
- Lobue, V., Nishida, T., Chiong, C., Deloache, J. S., & Haidt, J. (2011). When Getting Something Good is Bad: Even Three-year-olds React to Inequality. *Social Development*, *20*(1), 154–170. <https://doi.org/10.1111/j.1467-9507.2009.00560.x>
- Locey, M. L., Jones, B. A., & Rachlin, H. (2011). Real and hypothetical rewards. *Judgment and Decision Making*, *6*(6), 552–564.
<http://www.ncbi.nlm.nih.gov/pubmed/22582110>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3348706>
- Messick, D. M., & Schell, T. (1992). Evidence for an equality heuristic in social decision making. *Acta Psychologica*, *80*(1–3), 311–323. [https://doi.org/10.1016/0001-6918\(92\)90053-G](https://doi.org/10.1016/0001-6918(92)90053-G)
- Miller, D. T. (2001). Disrespect and the experience of injustice. *Annual Review of Psychology*, *52*. <https://doi.org/10.1146/annurev.psych.52.1.527>

- Munier, B., & Zaharia, C. (2002). High stakes and acceptance behavior in ultimatum bargaining: A contribution from an international experiment. *Theory and Decision*, 53(3), 187–207. <https://doi.org/10.1023/A:1022815832351>
- Norton, M. I., & Ariely, D. (2011). Building a better America-one wealth quintile at a time. *Perspectives on Psychological Science*, 6(1), 9–12. <https://doi.org/10.1177/1745691610393524>
- Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. In *Nature* (Vol. 437, Issue 7063). <https://doi.org/10.1038/nature04131>
- Pfeiffer, T., Rutte, C., Killingback, T., Taborsky, M., & Bonhoeffer, S. (2005). Evolution of cooperation by generalized reciprocity. *Proceedings of the Royal Society B: Biological Sciences*, 272(1568). <https://doi.org/10.1098/rspb.2004.2988>
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science*, 300(5626). <https://doi.org/10.1126/science.1082976>
- Schnedler, W. (2022). The broken chain: Evidence against emotionally driven upstream indirect reciprocity. *Games and Economic Behavior*, 136, 542–558. <https://doi.org/10.1016/j.geb.2022.10.008>
- Schnedler, W., & Stephan, N. L. (2020). Revisiting a Remedy Against Chains of Unkindness. *Schmalenbach Business Review*, 72(3). <https://doi.org/10.1007/s41464-020-00090-2>
- Sigmund, K., Fehr, E., & Nowak, M. A. (2002). The economics of fair play. *Scientific American*, 286(1). <https://doi.org/10.1038/scientificamerican0102-82>
- Stanca, L. (2009). Measuring indirect reciprocity: Whose back do we scratch? *Journal of Economic Psychology*, 30(2), 190–202. <https://doi.org/10.1016/j.joep.2008.07.010>

- Strang, S., Grote, X., Kuss, K., Park, S. Q., & Weber, B. (2016). Generalized Negative Reciprocity in the Dictator Game – How to Interrupt the Chain of Unfairness. *Scientific Reports*, 6(1), 22316. <https://doi.org/10.1038/srep22316>
- Thaler, R. H., & Camerer, C. F. (1995). Ultimatums, Dictators and Manners. *Journal of Economic Perspectives* 9(2):209-19.
- Verona, E., & Sullivan, E. A. (2008). Emotional Catharsis and Aggression Revisited: Heart Rate Reduction Following Aggressive Responding. *Emotion*, 8(3), 331–340. <https://doi.org/10.1037/1528-3542.8.3.331>
- Vidmar, N. J. (2005). Retribution and Revenge. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.224754>
- Wu, Y., Zang, Y., Yuan, B., & Tian, X. (2015). Neural correlates of decision making after unfair treatment. *Frontiers in Human Neuroscience*, 9, 123. <https://doi.org/10.3389/fnhum.2015.00123>
- Yamagishi, T., Horita, Y., Mifune, N., Hashimoto, H., Li, Y., Shinada, M., Miura, A., Inukai, K., Takagishi, H., & Simunovic, D. (2012). Rejection of unfair offers in the ultimatum game is no evidence of strong reciprocity. *Proceedings of the National Academy of Sciences of the United States of America*, 109(50). <https://doi.org/10.1073/pnas.1212126109>