Running head: OSTRACISM AND SANCTIONING

Who to Punish?

How observers sanction norm-violating behavior in ostracism situations

Selma C. Rudert

University of Koblenz-Landau

Simone Ruf and Rainer Greifeneder

University of Basel

Words: 10,893

Abstract

When observing ostracism, individuals can either side with the target or the sources of ostracism. Here we demonstrate that side-taking depends on whether the target previously acted in adherence to or in violation of perceived social norms. In four studies, a target behaved either norm-consistently or violated a social norm, and was subsequently either excluded by the sources or was not. Next, participants could sanction the behavior of the observed persons by refraining to assign money (Studies 1 and 2), or by subtracting money from a bonus (Studies 3 and 4). Observers assigned less money to the sources when these excluded a norm-consistent target. However, when the target had violated a social norm before, participants assigned less money to the target instead. These results have far-reaching implications because the (in)actions of neutral individuals can legitimize the sources' behavior, or help a target under attack.

Keywords: ostracism, punishment, observers, social norms

144 words

Ostracism and social exclusion are ubiquitous phenomena that oftentimes involve observers witnessing the ostracism situation. Many contributions have shown that observers empathize with the ostracized *target* and try to help or compensate him/her (see Wesselmann, Williams, & Hales, 2013), as well as punish the ostracizing *sources* for their inadequate behavior (Güroğlu, Will, & Klapwijk, 2013; Over & Uskul, 2016; Will, Crone, van den Bos, & Güroğlu, 2013). However, sometimes observers choose to do nothing, or even turn against the target themselves (Wesselmann, Wirth, Pryor, Reeder, & Williams, 2013). Here, we introduce norm-consistency of the target's behavior as one key moderator that explains whether observers perceive ostracism as unfair behavior of the sources that needs to be sanctioned, or a legitimate means to punish the target. More specifically, we suggest that if the target behaved consistently with social norms, participants will perceive ostracism as unacceptable, unfair behavior that needs to be sanctioned. However, when the target has violated social norms, ostracism is likely to be perceived as a legitimate means to punish the target for the deviation and, thus, constitutes a more acceptable behavior.

Understanding observers' side-taking is critical, since plenty of research has shown that ostracism is highly painful and threatening for its targets (for an overview, see Blackhart, Nelson, Knowles, & Baumeister, 2009; Hartgerink, van Beest, Wicherts, & Williams, 2015; Williams, 2009). Far less work has focused on the role of observers, who have the potential to make a difference by providing support and acknowledgement for ostracized targets, be it in an educational context, at work, or at home (Rudert & Greifeneder, 2019; Rudert, Hales, Greifeneder, & Williams, 2017). Yet, observers will not always side with the target and under certain conditions they may even join in ostracizing the target (Rudert, Reutner, Greifeneder, & Walker, 2017; Wesselmann, Wirth, et al., 2013; Wesselmann, Wirth, Pryor, Reeder, & Williams, 2015). Thus, from a theoretical as well as a practical perspective, it is critical to understand how observers make moral judgements and how these judgements affect subsequent behavior.

Ostracism and Social Norms

Individuals' behavior in social contexts is often driven and constrained by social norms that act as mental representations of appropriate situational behavior in a specific context (Aarts & Dijksterhuis, 2003; Cialdini & Trost, 1998). Social norms typically entail (a) a collective agreement on how individuals ought to act in a situation, (b) an expectation that they will act in accordance with the norm, as well as (c) the probability that violations of the social norm will be punished by others (Cialdini & Trost, 1998; Gibbs, 1965). While social norms may sometimes take the form of explicit rules, oftentimes they develop implicitly out of individuals' day-to-day interactions.

As social beings who thrive on cooperation, the default social norm in many human societies is one of inclusion. Thus, individuals have the normative expectation to be included in groups or ongoing activities, such as ball games or social activities (Rudert & Greifeneder, 2016; Wesselmann, Butler, Williams, & Pickett, 2010; Wesselmann, Wirth, & Bernstein, 2017). Ostracism without any apparent reason violates this default inclusion norm and is generally perceived as unfair or hostile (Rudert & Greifeneder, 2016; Tuscherer et al., 2015). For this reason, it is not surprising that many studies show that observers empathize with the target (Masten, Morelli, & Eisenberger, 2011; Wesselmann, Bagg, & Williams, 2009) and punish the sources, for instance, by decreasing their payoffs or choosing unequal distributions in a dictator game (Güroğlu et al., 2013; Over & Uskul, 2016; Will et al., 2013). Observers even choose to punish at their own expense, that is, when punishment has financial costs (Güroğlu et al., 2013; Will et al., 2013).

But groups and societies know social norms other than the inclusion norm, too, such as honesty, trustworthiness, or contributions to collective goods, and upholding these other social norms may prove vital for a group's or society's survival. One means of upholding social norms is ostracizing those who are deviant (Ditrich & Sassenberg, 2016; Stamkou, van Kleef, Homan, & Galinsky, 2016). In this case, ostracism serves a greater good and should be

accepted as legitimate by observers (punitive ostracism; Williams, Shore, & Grahe, 1998). As a consequence, observers should side with the sources and neither compensate the target nor sanction the sources for their ostracizing behavior. Consistent with this notion, Wesselmann and colleagues (2013) showed that individuals no longer compensated an ostracized target when s/he threw the ball in a ball-tossing game more slowly than everyone else.

From this norm perspective, it is critical to understand *which* social norms are at play when individuals observe ostracism. Being the default, the inclusion norm is primary (Pryor, Reeder, Wesselmann, Williams, & Wirth, 2013; Rudert & Greifeneder, 2016; Wesselmann, Wirth, et al., 2013). As has been demonstrated in previous research (Rudert, Sutter, Corrodi, & Greifeneder, 2018), when there is no apparent reason for ostracizing the target, observers attribute ostracism to malicious motives of the sources and consequently evaluate the sources negatively compared to the target, such as being less interested in cooperating with them, as well as reporting more anger and less sympathy towards the sources. In contrast, when observers perceive ostracism as a justified punishment, they are more likely to negatively evaluate the target compared to the sources. Here, we predict that these moral evaluations will be associated with individuals' behavior, so that observers who do not feel that the target has violated a social norm choose to side with the target and sanction the sources for their exclusionary behavior. In contrast, when observers know that a target has violated a social norm, observers are more likely to negatively evaluate and sanction the target compared to the sources in order to help protect the respective norm.

We put these considerations to the test in four studies, which investigated how observers sanction the behavior of both targets and sources of ostracism. We define sanctions as both disadvantaging one or more persons when distributing a resource as well as actively punishing by subtracting a resource (here: money) from others. The reported studies have been approved by the Institutional Review Board of the Faculty of Psychology, University of Basel, and conform to recognized standards written in the Declaration of Helsinki. All

participants provided informed consent at the beginning of the study. Data was stored in the departmental repository and will be made available upon request to rainer.greifeneder@unibas.ch.

Study 1

Method

      Participants and design. Participants were recruited online via several online platforms and mailing lists and had the option to participate in a lottery for book vouchers. We calculated the sample size to detect medium-sized effects ($f = .25$, power $= .80$, required $n = 128$). One hundred and twenty-five participants finished the questionnaire, however, four persons indicated that their data should not be used for analysis. Thus, the final sample consisted of 121 participants (91 females, $M_{age} = 25.69$, $SD = 8.20$) who were randomly assigned to a 2 (ostracism: exclusion vs. inclusion) x 2 (target behavior: norm-consistent vs. norm-violating) factorial design. In addition, sanctions of target and sources were assessed as repeated measures.

      Materials and procedure. Participants were told that they should evaluate a social interaction sequence within a group. They were presented with a group of three psychology students (Karin, Leonie, and Sandra) who took a statistics exam. Because the names were randomized to agents, we refer to the students and their roles as A, B, and C in what follows. Students B and C were unhappy, because they had achieved the lowest possible grade. In order to keep the conditions as similar as possible, in both the norm-consistent as well as the norm-violating condition, Student A lied to B and C about her grade. However, in the *norm-consistent* condition, Student A told a white lie (thus avoiding appearing superior to B and C) and claimed to have achieved the lowest possible grade as well, despite actually having achieved the highest possible grade. In the *norm-violating* condition, Student A is ostensibly lying to make herself appear better than B and C and claimed to have achieved the highest

possible grade, despite actually having achieved the lowest possible grade. In both conditions, Students B and C eventually find out about the lie. The story continued with the three students playing a ball game at a university sports event together. To illustrate the game, participants then watched a game of Cyberball from an observer's perspective in which Student A was either included in the game or excluded, that is, Students B and C did not throw the ball to her after a short initial period (for more details about Cyberball, see Williams, Cheung, & Choi, 2000).

Next, participants were asked to imagine that the three persons had found money on the street independently from each other. Participants were asked to decide how much money each person should find (0 - 5 Swiss Francs [CHF] per person, in increments of CHF 1; note that what was allocated to one was not deducted from the others). As money distribution for the two sources was highly correlated (Spearman's ρ = .94), values were averaged for the analysis. For exploratory reasons, participants were further asked how responsible each of the three persons was for the course of the story *(1 = not at all, 7 = very strongly)*; see online supplement for the results[1]. In addition, we assessed manipulation checks and control questions: Participants were asked to recall which grade each of the three students received. To measure whether the target's behavior was indeed perceived as norm-consistent or norm-violating, participants evaluated the fact that Student A was not telling the truth about her grade on three 7-point semantic differentials (*unacceptable – acceptable, undesirable – desirable, morally wrong – morally good*). Finally, the success of the ostracism manipulation was probed by means of the following questions: *"How much did the three persons participate in the ball game?" (1 = not at all, 7 = very much)* and how much each of the three persons was excluded and ignored (7-point semantic differential, *excluded – included, ignored – acknowledged*). After providing demographic information, participants were thanked, debriefed, and offered participation in the lottery.

Results

Manipulation checks. Eighty-six percent of participants correctly recalled each of the students' grades.[2] To test the successful manipulation of the target behavior, we ran a 2 (ostracism: exclusion vs. inclusion) x 3 (target behavior: norm-consistent vs. norm-violating) mixed MANOVA on acceptability, moral goodness, and desirability of the lie. There was only a significant effect of the target behavior, Wilks' λ = .804, $F(3, 114) = 9.29$, $p < .001$, η^2 = .20, while both the effect of ostracism and the interaction were non-significant, Wilks' λ = .979, $F(3, 114) = .82$, $p = .488$, $\eta^2 = .02$ and Wilks' λ = .993, $F(3, 114) = .27$, $p = .847$, η^2 = .01. Compared to the norm-violating lie, participants rated the norm-consistent (white) lie as more acceptable, $F(1, 116) = 8.25$, $p = .005$, $\eta^2 = .07$ and morally good, , $F(1, 116) = 18.94$, $p < .001$, $\eta^2 = .14$, though not as more desirable, $F(1, 116) = 1.23$, $p = .269$, $\eta^2 = .01$.

To test the successful manipulation of ostracism, we ran a 2 (person: target vs. sources) x 2 (ostracism: exclusion vs. inclusion) x 3 (target behavior: norm-consistent vs. norm-violating) mixed MANOVA with repeated measures on the person factor on perceptions of participation, exclusion, and inclusion. There were significant main effects of the person, Wilks' λ = .317, $F(3, 112) = 80.47$, $p < .001$, $\eta^2 = .68$, and ostracism, Wilks' λ = .460, $F(3, 112) = 12.60$, $p < .001$, $\eta^2 = .25$, that were qualified by a person x ostracism interaction, Wilks' λ = .386, $F(3, 112) = 59.29$, $p < .001$, $\eta^2 = .61$. Neither the main effect of the target behavior nor any of the interactions with target behavior were significant, largest F: Wilks' λ = .972, $F(3, 112) = 1.06$, $p = .368$, $\eta^2 = .03$. Breaking the interaction down by simple effects, compared to the inclusion condition, in the exclusion condition participants indicated that the *target* participated less, $F(1, 114) = 117.26$, $p < .001$, $\eta^2 = .51$, was more excluded, $F(1, 114) = 146.43$, $p < .001$, $\eta^2 = .56$, and ignored, $F(1, 114) = 65.80$, $p < .001$, $\eta^2 = .37$. In contrast, participants felt that compared to the inclusion group, excluding *sources* participated more, $F(1, 114) = 54.04$, $p < .001$, $\eta^2 = .32$, were less excluded, $F(1, 114) = 18.40$, $p < .001$, $\eta^2 = .14$, and less ignored, $F(1, 114) = 19.41$, $p < .001$, $\eta^2 = .15$. See *Table 1* for the descriptive statistics.

Dependent variables. We ran a 2 (person: target vs. sources) x 2 (ostracism: exclusion vs. inclusion) x 2 (target behavior: norm-consistent vs. norm-violating) mixed ANOVA with repeated measures on the person factor on the distributed amount of money. None of the main effects were significant (person: $F(1, 117) = 1.24$, $p = .269$, $\eta^2 = .01$; ostracism: $F(1, 117) = 1.67$, $p = .199$, $\eta^2 = .01$; target behavior: $F(1, 117) = 0.90$, $p = .344$, $\eta^2 = .01$). The person x ostracism and the ostracism x target behavior interactions were also not significant, $F(1, 117) = 0.45$, $p = .506$, $\eta^2 = .00$ and $F(1, 117) = 0.38$, $p = .537$, $\eta^2 = .00$. However, the analysis revealed a significant person x target behavior two-way interaction, $F(1, 117) = 21.39$, $p < .001$, $\eta^2 = .16$, which was qualified by the hypothesized person x ostracism x target behavior three-way interaction, $F(1, 117) = 6.22$, $p = .014$, $\eta^2 = .05$.

To break down the three-way interaction, we ran the analysis separately for the two target behavior conditions: When the target violated the social norm, there was only a main effect of the person, showing that the target was sanctioned more than the sources, $F(1, 55) = 15.75$, $p < .001$, $\eta^2 = .22$. Ostracism and the interaction had no significant effects, $F(1, 55) = 0.21$, $p = .650$, $\eta^2 = .00$, and $F(1, 55) = 1.60$, $p = .211$, $\eta^2 = .03$ However, when the target acted consistently with the social norm, there was a significant effect of the person, $F(1, 62) = 6.48$, $p = .013$, $\eta^2 = .10$, that was qualified by the person x ostracism two-way interaction, $F(1, 62) = 5.24$, $p = .025$, $\eta^2 = .08$. There was no main effect of ostracism, $F(1, 62) = 1.99$, $p = .163$, $\eta^2 = .03$. Breaking down the interaction via simple main effects, we found that when the sources excluded the target, they were sanctioned more than the target, $F(1, 62) = 12.47$, $p < .001$, $\eta^2 = .17$. There was no significant effect in the inclusion condition, $F(1, 62) = 0.03$, $p = .862$, $\eta^2 = .00$. See *Figure 1* and *Table 1* for the descriptive statistics.

Discussion

In line with our hypotheses, participants sanctioned ostracism of a target without a proper reason by allocating less money to the sources. However, when the target had previously violated a social norm, observers sanctioned the target's behavior instead.

One caveat of the study is that perceived norm-consistency was likely only relative between conditions, as descriptively, in all conditions Student A's behavior was rated below the scale midpoint of 4 on acceptability, moral goodness, and desirability of the lie. Perhaps participants perceived lying as always being wrong, no matter whether it was a white lie or not. We address this potential limitation in Study 2.

Study 2

Because investigating behaviors that are perceived as norm-consistent or violating social norms is central to our hypothesis, in Study 2, we chose a scenario in which the target acted fully appropriately in the norm-consistent condition. Moreover, we split the norm-violating condition into a "weak norm-violation" and a "strong norm-violation," to investigate differences in the appropriateness of exclusion as a punishment. Finally, we exchanged the inclusion group with a control group that received no information about the sources' actions.

Method

Participants and design. Participants were recruited online via several online platforms and mailing lists and had the option to participate in a lottery for book coupons. We calculated the sample size so as to detect medium-sized effects ($f = .25$, power $= .80$, required $n = 158$). One hundred and seventy-two participants finished the questionnaire. However, three persons indicated that their data should not be used for analysis. Thus, the final sample consisted of 169 participants (126 females, $M_{age} = 25.64$, $SD = 8.05$) who were randomly assigned to a 2 (ostracism: exclusion vs. control) x 3 (target behavior: norm-consistent vs. weakly norm-violating vs. strongly norm-violating) factorial design. Sanctions of target and sources were assessed as repeated measures.

Materials and procedure. Participants were told that they should evaluate a social interaction. They were presented with a story about a group of three female students (Karin, Leonie, and Sandra) who were friends attending a party together. As in Study 1, the names were randomized to agents, and we refer to the students and their roles as A, B, and C in what

follows. Student B brought her new boyfriend with her to the party, and this boyfriend later starts to flirt with Student A. We manipulated Student A's reaction: In the norm-consistent condition, Student A ended the conversation with the boyfriend immediately when he started flirting, and walked away. In the weak norm-violation-condition, she flirted back at first, but ended the conversation when he got too close. In the strong norm-violating condition, she started kissing Student B's boyfriend. In all three conditions, Students B and C noticed what was going on. The three descriptions of the target's behavior were pre-tested with 75 participants via the mailing list of the online psychology Platform www.forschung-erleben.de (53 females, M_{age} = 27.39, SD = 6.36). Note that in the pre-test, three versions of the weak norm-violation description were tested, but here we only report the results of the version that was also used in the main study. Pretest-participants evaluated Student A's behavior on four 7-point semantic differentials (*unacceptable – acceptable, undesirable – desirable, morally wrong – morally good, incorrect - correct),* that were aggregated to one moral judgement score, Cronbach's α = .94. An ANOVA of target behavior on moral judgement showed significant differences between the conditions, $F(2, 41)$ = 20.34, $p <$ .001, η^2 = .50. The norm-consistent target behavior was rated as morally superior to the weak violation (M = 6.08, SD = 1.58 vs. M = 4.55, SD = 1.59), $t(28)$ = 2.65, p = .013, d = 0.97, and both the norm-consistent behavior and the weak violation were rated as morally superior to the strong norm-violation (M = 2.70, SD = .92), $t(26)$ = 6.94, $p <$ .001, d = 2.62 and $t(28)$ = 3.83, p = .001, d = 1.40. The strong norm-violation and the norm-consistent behavior were both rated as significantly different from the scale-midpoint in the respective directions, $t(13)$ = -5.30, $p <$ .001, d = 1.41 and $t(13)$ = 4.94, $p <$ .001, d = 1.31.

As in Study 1, in the Exclusion condition, the story continued with the three students playing a ball game at a university sports event together the next day. Participants then watched a game of Cyberball in which Student A was excluded from the game. In the Control condition, there was no ball game and participants directly answered the dependent variables.

The dependent variable was the same as in Study 1 (money found on the street; 0 - 5 Swiss

Francs [CHF], in increments of 1 CHF). Values for the two sources were averaged

(Spearman's ρ = .93). As in Study 1, participants were also asked about the level of

responsibility each of the three students had over the course of the events; see online

supplement. Participants also evaluated the morality of Student A's behavior on the same

items as in the pretest, Cronbach's α = .92. The morality of the boyfriend's behavior was also

assessed for exploratory reasons unrelated to the present research question, thus, no respective

analyses are reported. Participants were further asked to identify the name of Student A, who

was being flirted with, and the name of Student B, who brought her boyfriend to the party. In

addition, participants in the exclusion condition evaluated the students' participation in the

Cyberball game and how much each of the students was excluded and ignored (see Study 1).

After providing demographic information, participants were thanked, debriefed, and could

participate in the lottery.

Results

       Manipulation checks. Ninety-one percent of participants correctly recalled the names

of Students A and B. The target's behavior significantly affected participants' moral

judgment, $F(2, 162) = 42.84$, $p < .001$, $\eta^2 = .35$. Neither ostracism nor the interaction had an

effect, $F(1, 162) = 0.09$, $p = .769$, $\eta^2 = .00$ and $F(2, 162) = 0.62$, $p = .538$, $\eta^2 = .01$. Compared

to the norm-consistent behavior, both the weak and the strong norm-violation were rated as

more morally wrong, $t(1,162) = 7.64$, $p < .001$, $d = 1.20$ and $t(1,162) = 8.42$, $p < .001$, $d =$

1.32; however, there was no significant difference between the weak and the strong norm-

violation, $t(1,162) = 0.67$, $p = .503$, $d = 0.11$.

       As the ostracism manipulation check was only presented to participants in the

ostracism but not in the control condition, the control and the ostracism condition cannot be

compared. As an equivalent, we thus compared the targets and the sources. A 2 x (person:

target vs. sources) x 3 (target behavior: norm-consistent vs. weak norm-violation vs. strong

norm-violation) MANOVA with repeated measures on the person factor on participation, exclusion, and ignoring showed only an effect of the person, Wilks' λ = .066, $F(3, 66)$ = 310.86, $p < .001$, η^2 = .68, there was no effect of target behavior, Wilks' λ = .963, $F(6, 132)$ = 0.418, p =.868, η^2 = .02, or an interaction, Wilks' λ = .896, $F(6, 132)$ = 1.24, p = .291, η^2 = .05. Participants in the exclusion condition indicated that the target participated less than the sources, $F(1,68)$ = 899.57, $p < .001$, η^2 = .93, was more excluded, $F(1,68)$ = 807.56, $p < .001$, η^2 = .92, and ignored, $F(1,68)$ = 504.56, $p < .001$, η^2 = .88. See *Table 2* for the descriptive statistics.

Dependent variables. We ran a 2 x (person: target vs. sources) 2 (ostracism: exclusion vs. control) x 3 (target behavior: norm-consistent vs. weak norm-violation vs. strong norm-violation) mixed ANOVA with repeated measures on the person factor on the amount of money that participants distributed. The analysis revealed main effects for ostracism, $F(1, 161)$ = 4.48, p = .036, η^2 =.03, and for the target behavior, $F(2, 161)$ = 3.92, p = .022, η^2 =.05, which were both qualified by a significant person x target behavior two-way interaction, $F(2, 161)$ = 14.96, $p < .001$, η^2 =.16, a significant person x ostracism two-way interaction, $F(1, 161)$ = 14.54, $p < .001$, η^2 =.08, and a significant target behavior x ostracism two-way interaction, $F(2, 161)$ = 6.25, p = .002, η^2 =.07. The main effect of the person and the three-way interaction were not significant, $F(1, 161)$ = 1.20, p = .370, η^2 =.01 and $F(2, 161)$ = 1.20, p = .304, η^2 =.02.

Although the three-way interaction was not significant, we offer an independent analysis of the results separately for target behavior condition so as to ensure comparability with the results' sections of the other studies: As in Study 1, when the target violated the social norm, there was only a main effect of the person, showing that the target was sanctioned more than the sources, $F(1, 56)$ = 18.02, $p < .001$, η^2 =.24. Neither the effect of ostracism nor the interaction was significant, $F(1, 56)$ = 1.99, p = .164, η^2 =.03 and $F(1, 56)$ = 1.15, p = .288, η^2 =.02.

When the target committed a weak norm violation, there was a significant effect of ostracism, $F(1, 52) = 4.58$, $p = .037$, $\eta^2 = .08$, that was qualified by the significant person x ostracism two-way interaction, $F(1, 52) = 4.59$, $p = .037$, $\eta^2 = .08$. The main effect of the person was not significant, $F(1, 52) = 0.05$, $p = .831$, $\eta^2 = .00$. Breaking down the interaction via simple main effects, we found that targets and sources were not treated differently in the exclusion group, $F(1, 52) = 1.62$, $p = .209$, $\eta^2 = .03$. In the control group, however, there was a non-significant trend that the target was sanctioned more than the sources, $F(1, 52) = 3.26$ $p = .077$, $\eta^2 = .06$.

When the target acted consistently with the social norm, there was a significant effect of the person, $F(1, 53) = 11.89$, $p = .001$, $\eta^2 = .18$, and of ostracism, $F(1, 53) = 10.56$, $p = .002$, $\eta^2 = .17$, that was qualified by the person x ostracism two-way interaction, $F(1, 53) = 13.12$, $p = .001$, $\eta^2 = .20$. Breaking down the interaction via simple main effects, we found that when the sources excluded the target, they were sanctioned more than the target, $F(1,53) = 20.83$, $p < .001$, $\eta^2 = .28$. In the control group, targets and sources were not treated differently, $F(1,53) = 0.02$, $p = .891$, $\eta^2 = .00$. See *Figure 2* and *Table 2* for the descriptive statistics.

Discussion

Study 2 replicates the findings of Study 1: Participants assigned less money to the sources for excluding a norm-consistent target, but when the target had violated social norms, participants sanctioned the target's behavior instead. Interestingly, observers' moral judgement was more extreme than the sanctioning behavior they ultimately displayed: While participants perceived both the weak and the strong norm-violating behavior of the target as morally wrong, they only accepted the target's exclusion and assigned less money to her than to the sources when she had committed a strong but not a weak norm-violation.

Study 3

Studies 1 and 2 relied on scenarios, such that participants knew that they were not actually assigning money to others. While money allocation is an often-used variable to

measure punishment in ostracism contexts (Güroğlu et al., 2013; Over & Uskul, 2016; Will et al., 2013), one could argue that fictional money distribution is not a fitting sanction for the norm violations in Studies 1 and 2. Moreover, as the money distribution was fictional, it might rather measure participants' judgment of deservingness, which may or may not translate into participants actively aiming to restore justice by punishing the sources or the target. To investigate this question, Study 3 was designed in such a way that participants assumed they were actually punishing a third person and provided a better fit between the norm violation and the punishment. In addition, we tested the assumption that the effect of norm-conformity and ostracism on the sanctioning of targets and sources is mediated via the evaluation of targets and sources.

Method

Participants and design. Participants were recruited via Prolific Academic (US Americans only) for a payment of £1. We calculated the minimal sample size to detect medium-sized effects ($f = .25$, power $= .90$, required $n = 171$). One hundred and ninety-nine participants finished the questionnaire (100 females, 1 person who did not assign a gender category), $M_{age} = 31.79$, $SD = 11.41$) and were randomly assigned to a 2 (ostracism: exclusion vs. inclusion) x 2 (target behavior: norm-consistent vs. norm-violating) factorial design. Punishment of target and sources was assessed as repeated measures.

Materials and procedure. Participants were told that they should evaluate a social interaction. They were presented with material from an alleged previous study in which participants (henceforth called "players") had to work together as a team on several tasks. In that study, the alleged players had played a cooperation game called "the fishpond," a dilemma of the commons (Spada & Opwis, 1985), in which players can draw fish from a common pond over three rounds. In the norm-consistent condition, all players chose to play cooperatively, by taking only one fish in the first two rounds. Thus, the fish population could recover and everyone ended up with four fish at the end of round three. In the norm-violating

condition, the target acted greedily and took the maximum amount of fish (two) out of the pond each round. S/he ended up with six fish while the other two players only received two fish each. After the fishpond game, the players allegedly worked together on a follow-up task. For this task, they could decide whether they would like to work with both participants or decline to work with one or even both of the others. In the exclusion condition, the two sources both declined to work with the target; in the inclusion condition, every player chose to work with both co-players.

After observing the interaction within the alleged study, the actual participants were told that each of the players would receive a £1 bonus payment for each fish that s/he caught during the fishpond game. Participants were further told that they had the possibility to reduce the bonus payments that the three players would receive. The final bonus payment received by the game players would allegedly be determined by the number of fish the players had caught minus the average penalty given by the other participants.

As ratings of the two sources were highly correlated (Spearman's ρ = .95 - .98), ratings were averaged for the dependent variable and the mediator. To assess punishment, participants had the option to subtract money up to a maximum of £2 in increments of £0.25 from each player's bonus. Moreover, to assess how participants evaluated the players, we asked participants three items per player measuring their cognitive and emotional responses towards the other players. Specifically, we asked how much they would like to work together on a cooperative task with each of the players *(1 = not at all, 7 = very much)* and how they felt about the other persons and their behavior: *"I feel angry about the behavior of [Player's Name]"* and *"I can sympathize with [Player's Name]" (1 = not at all, 7 = very much),* for all scales, see Rudert and colleagues (2018). An exploratory factor analysis extracted one factor that explains 77% of the variance for the target and 72% for the sources; all factor loadings were > .80. The three evaluation measures were thus aggregated in a composite evaluation score (Target: α = .85; Sources: α = .80).

We further assessed ascribed responsibility (see online supplement for both the item descriptions and results). As a manipulation check, participants were asked which player(s) had the most fish at the end of the fishpond game and which player would have to work alone in the subsequent group task. We further asked how realistic the situation was; for item descriptions and results, see online supplement. Moreover, we added several attention checks throughout the study and informed participants upfront about their occurrence, which was meant to ensure that participants paid attention. These variables were not analyzed.

Results

Manipulation checks. 95% of all participants correctly answered which player caught the most fish and 94% correctly recalled which player had to work alone in the group task. In total, 90% of the participants answered both manipulation checks correctly.

Dependent variables. *Punishment.* We ran a 2 (person: target vs. sources) x 2 (ostracism: exclusion vs. inclusion) x 2 (target behavior: norm-consistent vs. norm-violating) mixed ANOVA with repeated measures on the person factor on the amount of money that participants subtracted. The analysis revealed an effect for the person, $F(1, 195) = 79.30$, $p < .001$, $\eta^2 = .29$ and for the target behavior, $F(1, 195) = 37.81$, $p < .001$, $\eta^2 = .16$, which were both qualified by a person x target behavior two-way interaction, $F(1, 195) = 118.61$, $p < .001$, $\eta^2 = .38$, and a person x ostracism x target behavior three-way interaction, $F(1, 195) = 4.85$, $p = .029$, $\eta^2 = .02$. Neither the effect of ostracism, $F(1, 195) = 0.94$, $p = .333$, $\eta^2 = .01$, nor the two-way interactions of ostracism x person and ostracism x conformity were significant, $F(1, 195) = 0.01$, $p = .946$, $\eta^2 = .00$ and $F(1, 195) = 0.02$, $p = .900$, $\eta^2 = .00$. See *Figure 3* and *Table 3* for the descriptive statistics.

To break down the three-way interaction, we ran the analysis separately for the target behavior conditions: When the target violated the social norm, there was only a main effect of the person, showing that the target was sanctioned more than the sources, $F(1, 100) = 122.05$,

$p < .001$, $\eta^2 = .55$. Ostracism and the interaction had no significant effects, $F(1, 100) = 0.57$, $p = .453$, $\eta^2 = .01$ and $F(1, 100) = 1.60$, $p = .208$, $\eta^2 = .02$.

When the target acted in line with the social norm, ostracism had no significant main effect, $F(1, 95) = 0.38$, $p = .538$, $\eta^2 = .00$, but there was a significant main effect of the person, $F(1, 95) = 6.06$, $p = .016$, $\eta^2 = .06$, that was qualified by a significant person x ostracism two-way interaction, $F(1, 95) = 7.00$, $p = .010$, $\eta^2 = .07$. Looking at the simple main effects, we found that when the sources excluded the target, they were sanctioned more than the target, $F(1, 95) = 13.75$, $p < .001$, $\eta^2 = .13$. There was no significant difference between target and sources in the inclusion group, $F(1, 95) = 0.02$, $p = .899$, $\eta^2 = .00$.

Evaluation. An ANOVA on evaluation revealed a significant main effect for all independent variables (person: $F(1, 195) = 43.77$, $p < .001$, $\eta^2 = .18$, ostracism: $F(1, 195) = 20.72$, $p < .001$, $\eta^2 = .10$, target behavior: $F(1, 195) = 46.20$, $p < .001$, $\eta^2 = .19$) as well as two two-way interactions (person x target behavior: $F(1, 195) = 194.40$, $p < .001$, $\eta^2 = .50$, ostracism x target behavior: $F(1, 195) = 8.85$, $p = .003$, $\eta^2 = .04$, that were qualified by a significant three-way interaction, $F(1, 195) = 27.46$, $p < .001$, $\eta^2 = .12$. The person x ostracism interaction did not reach significance, $F(1, 195) = 3.62$, $p = .059$, $\eta^2 = .02$,

Similar to participants' punishment behavior, when the target had violated the social norm, there was a main effect of the person insofar that participants evaluated the target worse, $F(1, 100) = 151.52$, $p < .001$, $\eta^2 = .60$. The main effect of ostracism was not significant, $F(1, 100) = 1.84$, $p = .178$, $\eta^2 = .02$, but there was a significant interaction, $F(1, 100) = 3.99$, $p = .048$, $\eta^2 = .04$. Looking at the simple main effects, the target was evaluated worse than the sources when s/he was both included or excluded, but the effect sizes indicate that the effect was larger in the exclusion compared to the inclusion group, $F(1, 100) = 100.39$, $p < .001$, $\eta^2 = .50$ and $F(1, 100) = 54.22$, $p < .001$, $\eta^2 = .35$.

When the target had acted norm-consistently, there were significant main effects of the person, $F(1, 95) = 48.00$, $p < .001$, $\eta^2 = .34$ and of ostracism, $F(1, 95) = 20.89$, $p < .001$, η^2

=.18, that were qualified by a significant person x ostracism two-way interaction, $F(1, 95) =$ 18.05, $p < .001$, $\eta^2 = .32$. Looking at the simple main effects, we found that when the sources excluded the target, participants evaluated the sources worse than the target, $F(1, 95) = 98.68$, $p < .001$, $\eta^2 = .51$. When the sources included the target, evaluation of the target and the sources did not differ, $F(1, 95) = 0.01$, $p = .905$, $\eta^2 = .00$.

Mediation analyses. We ran two mediation models with the SPSS PROCESS macro (Hayes, 2013), using 5,000 bootstrap estimates: (a) a mediation model for the effect of conformity on punishment of the target via evaluation of the *target* and (b) a mediated moderation model for the effect of the ostracism x conformity interaction on punishment of the sources via evaluation of the *sources*.

As for model (a), there was a significant indirect effect of conformity on punishment of the target via evaluation of the target, $b_{indirect} = 2.62$, 95% CI [1.83; 2.46]. If the target violated the social norm, s/he was evaluated more negatively and this was associated with more money being subtracted from his/her bonus.

As for model (b), the effect of the ostracism x conformity interaction on punishment of the sources was significantly mediated via evaluation of the sources, $b_{indirect} = -1.07$, 95% CI = [-2.07; -.36]. If the sources excluded compared to included a norm-consistent target, the sources were evaluated more negatively and this was associated with more money being subtracted from their bonus, $b_{indirect} = .93$, 95% CI = [.29; 1.79]. If the target had violated the social norm, the indirect effect was not significant, $b_{indirect} = -.14$, 95% CI [-.45; .03].

Discussion

Study 3 replicates and further extends the findings of Studies 1 and 2: When the target had violated social norms before, observers punished the target. When the target had acted norm-consistently but was excluded by the sources, participants punished the sources. These effects were mediated via the evaluation of the target and the sources, respectively, suggesting

that observers dislike norm violations (violations of the norm of cooperation or inclusion) and

consequently devalue as well as punish the respective perpetrators.

Study 4

Study 4 was similar to Study 3 and served to address open alternative explanations.

First, in Study 3 it might have been unclear whether the effects were due to norms being

violated or due to the target committing some sort of interpersonal transgression against the

sources. Indeed, one could argue that taking more fish, knowing that the others will then

receive less, could be perceived as an interpersonal transgression. To investigate this potential

confound, we did two things: we ran a pre-test showing that there is a general injunctive norm

to behave cooperatively within the Fishpond Game (that is, individuals believe that one

should play cooperatively within the game). In addition, we changed the game in Study 4 so

that in none of the conditions did the sources suffer a direct disadvantage from the target's

behavior.

Second, another potential concern might have been that the game was not very

consequential for the observers. Specifically, in Study 3, the observing participants could

make a decision about punishment without having to bother about downstream consequences

for themselves. Yet in real life, observers may need to be careful if they decide to punish

someone, because if this course of action itself violates social norms, others may disagree and

observers might easily become the target of repercussions themselves. To simulate such a

consequential situation, Study 4's participants were incentivized to act in line with what

others perceive as acceptable. Specifically, participants were told that they would receive a

higher bonus payment if they acted in line with the average penalties that other participants in

the same social situation had assigned.

Third, in Study 3 we focused on the mediating role of evaluations, yet one may argue

that perceptions of fairness play an important role, too. Potential evidence stems from

research on third-party punishment showing that defection in cooperative games elicits

negative fairness judgements as well as punishment behavior in observers (Fehr &

Fischbacher, 2004). Moreover, fairness considerations play an important role in models of

observer's reactions to workplace injustice (Skarlicki, O'Reilly, & Kulik, 2015). Thus, in

Study 4 we additionally test both the role of evaluations as well as perceptions of fairness as

potential mediators of the effect of target behavior and ostracism on sanctioning by the

observers.

We originally predicted and pre-registered a serial mediation model, assuming that

that target's behavior would affect observer's fairness considerations, which would then affect

evaluations of targets and sources, and ultimately participants' punishing behavior. However,

during the review process, the Editor asked us to perform a simultaneous mediation analysis

instead, given that the study design does not allow for causal conclusions regarding the

mediation. The model reported in what follows thus predicts that participants judge the

behavior of a norm-violating target (compared to a norm-consistently acting target) as unfair

and evaluate the target negatively, which should be associated with greater punishment of the

target. However, when the target behaves in a norm-consistent manner and the sources

exclude her/him nonetheless, the model predicts that observers perceive the sources' behavior

as unfair and evaluate the sources negatively, which should be associated with stronger

punishment of the sources. For results of the serial mediation model, see online supplement.

Pretest

For the pretest, we recruited 81 participants (37 females, M_{age} = 34.94, SD = 11.12) via

Prolific Academic for a payment of £0.50. They were presented with the rules of the Fishpond

game (without observing the game being played) and were then randomly assigned to two

norm conditions. Participants in the *injunctive norm condition* were asked what other

individuals would generally *expect* to be the right thing to do when playing the Fishpond

game. Participants in the *descriptive norm condition* were asked what individuals would

generally *do* when playing the Fishpond game. The response options were the same for both

groups (take zero/one/more than one fish per round). Results show that in the injunctive norm condition, 75% of the participants replied that each player should take one fish per round, and 25% replied that they should take more than one, $X^2(1, n = 40) = 10$, $p = .002$. In the descriptive norm condition, 58% of the participants replied that each player would take more than one fish per round and 42% replied that they would take only one, $X^2(1, n = 40) = 0.90$, $p = .343$. Thus, while participants were not necessarily convinced that others *would in fact* play cooperatively (a pessimistic descriptive social norm), there was a clear injunctive norm that they *should* do so.

Method

Participants and design. Participants were recruited via Prolific Academic (US-Americans only) for a payment of £0.90 plus a variable bonus of up to 30p. Since the effect size of the three-way interaction in Study 3 was small, following Cohen's conventions we calculated the sample size based on an effect size of $f = .14$, aiming for a power = .80. G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) calculates a required sample of 400. To account for possible data exclusions, we chose to oversample by 5% and collected data of 420 participants. The study was pre-registered on AsPredicted (https://aspredicted.org/4jt9s.pdf). Participants were randomly assigned to a 2 (ostracism: exclusion vs. inclusion) x 2 (target behavior: norm-consistent vs. norm-violating) factorial design. Punishment of target and sources was assessed as repeated measures. In total, 425 participants finished the questionnaire, however, seven persons indicated that their data should not be used, so that the final sample consisted of 418 persons (207 females, 1 person who did not assign a gender category), $M_{age} = 33.85$, $SD = 12.15$).

Materials and procedure. The procedure was the same as in Study 3, with the following changes: (1) In Study 4, the sources had not played the fishpond game with the target. Instead, participants watched the target playing the fishpond game with two persons, and then being assigned to a different group. Within that new group, all players could see how

many fish the respective others had caught in the previous fishpond game. By means of this procedure, the sources knew whether the target had violated the social norm, but had not incurred a personal disadvantage from a target behaving uncooperatively. (2) Participants could win a bonus of a maximum of 30p after the study. This bonus was linearly decreased the more participants' decisions deviated from what can be expected as the descriptive norm in this situation (we used the decisions from participants in Study 3 as the descriptive norm). This payout-structure was implemented to make the study more consequential for participants, mirroring the social dilemma that observers often find themselves in when considering whether to intervene: If observers' decisions regarding punishment are not considered acceptable by others, observers might suffer from negative (social) consequences. (3) We included the evaluation measures used in Study 3 (interest in cooperation, anger, and sympathy with each of the players) and aggregated them in a composite evaluation score (Target: $\alpha = .86$; Sources: $\alpha = .78$). In addition, we included a measure of fairness to test its role as an additional mediator. Specifically, participants were asked, *"How fairly did the different players act throughout the study"* and rated each player on a 7-point scale (*1= very unfair; 7 = very fair*). As ratings of the two sources were highly correlated (Spearman's ρ = .95 - .98), ratings were averaged for the dependent variable and both mediators.

Responsibility was not assessed in Study 4. We added another open-format manipulation check asking about the maximum amount of fish that individuals who acted cooperatively could catch during the Fishpond game.

Results

Manipulation checks. 85% of all participants correctly answered which player caught the most fish and 90% correctly recalled which player had to work alone in the group task. In total, 79% of the participants answered both checks correctly. The majority of participants (81 %) also correctly stated that a cooperative player would end up with a maximum of four fish in the end.

Dependent variables.

Punishment. We ran a 2 (person: target vs. sources) x 2 (ostracism: exclusion vs. inclusion) x 2 (target behavior: norm-consistent vs. norm-violating) mixed ANOVA with repeated measures on the person factor on the amount of money that participants subtracted. The analysis revealed no significant main effect of ostracism, $F(1, 414) = 3.77$, $p = .053$, $\eta^2 =.01$. However, there was a significant main effect for the person, $F(1, 414) = 212.26$, $p < .001$, $\eta^2 =.34$ and for the target behavior, $F(1, 414) = 121.39$, $p < .001$, $\eta^2 =.23$, which were both qualified by a person x target behavior two-way interaction, $F(1, 414) = 356.16$, $p < .001$, $\eta^2 =.46$, and the hypothesized person x ostracism x target behavior three-way interaction, $F(1, 414) = 13.47$, $p < .001$, $\eta^2 =.03$. The ostracism x person and the ostracism x target behavior interaction were both not significant, $F(1, 414) = 1.78$, $p = .183$, $\eta^2 =.00$, $F(1, 414) = 1.78$, $p = .183$, $\eta^2 =.00$ and $F(1, 414) = 2.75$, $p = .098$, $\eta^2 =.01$. See *Figure 4* and *Table 4* for the descriptive statistics.

To break down the three-way interaction, we ran the analysis separately for the target behavior conditions[3]: When the target violated the social norm, there was only a main effect of the person, showing that the target was sanctioned more than the sources, $F(1, 207) = 345.78$, $p < .001$, $\eta^2 =.63$. Ostracism and the person x ostracism interaction were not significant, $F(1, 207) = 0.04$, $p = .849$, $\eta^2 =.00$ and $F(1, 207) = 1.69$, $p =.196$, $\eta^2 =.01$. When the target acted in line with the social norm, there was a significant effect of the person, $F(1, 207) = 24.16$, $p < .001$, $\eta^2 =.10$, and of ostracism, $F(1, 207) = 7.25$, $p = .008$, $\eta^2 =.03$, that were qualified by a significant person x ostracism two-way interaction, $F(1, 207) = 32.68$, $p < .001$, $\eta^2 =.14$. Looking at the simple main effects, we found that when the sources excluded the target, they were sanctioned more than the target, $F(1, 207) = 56.25$, $p < .001$, $\eta^2 =.21$. When the sources included the target, there was no significant difference between sources and target, $F(1, 207) = 0.324$, $p = .570$, $\eta^2 =.00$.

Evaluation. An ANOVA on evaluation revealed a significant main effect for all independent variables (person: $F(1, 414) = 128.17$, $p < .001$, $\eta^2 = .24$, ostracism: $F(1, 414) = 33.07$, $p < .001$, $\eta^2 = .07$, target behavior: $F(1, 414) = 266.26$, $p < .001$, $\eta^2 = .39$), as well as three two-way interactions (person x ostracism: $F(1, 414) = 12.00$, $p = .001$, $\eta^2 = .03$, person x target behavior: $F(1, 414) = 594.27$, $p < .001$, $\eta^2 = .59$, ostracism x target behavior: $F(1, 414) = 64.76$, $p < .001$, $\eta^2 = .14$), that were qualified by a significant three-way interaction, $F(1, 414) = 72.69$, $p < .001$, $\eta^2 = .15$.

To break down the three-way interaction, we ran the analysis separately for the target behavior conditions: When the target violated the social norm, there was a main effect of the person, $F(1, 207) = 477.44$, $p < .001$, $\eta^2 = .70$. The effect of ostracism was not significant $F(1, 207) = 2.80$, $p = .096$, $\eta^2 = .01$, but there was a significant person x ostracism interaction, $F(1, 207) = 9.60$, $p = .002$, $\eta^2 = .04$. Looking at the simple main effects, in both the inclusion and the exclusion condition, the target was generally evaluated worse than the sources. When comparing the effect sizes, the effect was stronger in the exclusion condition, $F(1, 207) = 318.85$, $p < .001$, $\eta^2 = .61$ than in the inclusion condition, $F(1, 207) = 171.72$, $p < .001$, $\eta^2 = .45$.

When the target acted in line with the social norm, there was a significant effect of the person, $F(1, 207) = 128.07$, $p < .001$, $\eta^2 = .38$, and of ostracism, $F(1, 207) = 89.85$, $p < .001$, $\eta^2 = .30$, that were qualified by a significant person x ostracism two-way interaction, $F(1, 207) = 107.99$, $p < .001$, $\eta^2 = .34$. Looking at the simple main effects, we found that when the sources excluded the target, participants evaluated them worse compared to the target, $F(1, 207) = 234.51$, $p < .001$, $\eta^2 = .53$. There was no significant difference in the evaluation of target and sources following inclusion, $F(1, 207) = 0.43$, $p = .513$, $\eta^2 = .00$.

Fairness. An ANOVA on fairness revealed a significant main effect for all independent variables (person: $F(1, 414) = 192.15$, $p < .001$, $\eta^2 = .32$, ostracism: $F(1, 414) = 59.79$, $p < .001$, $\eta^2 = .13$, target behavior: $F(1, 414) = 392.03$, $p < .001$, $\eta^2 = .49$), as well as

three two-way interactions (person x ostracism: $F(1, 414) = 10.47$, $p = .001$, $\eta^2 =. 03$; person x target behavior: $F(1, 414) = 616.19$, $p < .001$, $\eta^2 =. 60$, ostracism x target behavior: $F(1, 414) = 38.87$, $p < .001$, $\eta^2 =. 09$), that were qualified by a significant three-way interaction, $F(1, 414) = 61.31$, $p < .001$, $\eta^2 =.13$.

To break down the three-way interaction, we ran the analysis separately for the target behavior conditions: When the target violated the social norm, there was no significant effect of ostracism, $F(1, 207) = 1.25$, $p = .264$, $\eta^2 =.01$, however, there was a significant effect of the person, $F(1, 207) = 533.09$, $p < .001$, $\eta^2 =.72$, as well as a person x ostracism interaction, $F(1, 207) = 7.52$, $p = .007$, $\eta^2 =.03$. Looking at the simple main effects, the target was always rated less fair than the sources. Comparing the effect sizes, however, the effect was stronger when s/he had been excluded, $F(1, 207) = 341.79$, $p < .001$, $\eta^2 =.62$, compared to included, $F(1, 207) = 202.16$, $p < .001$, $\eta^2 =.50$.

When the target acted in line with the social norm, there was a significant main effect of the person, $F(1, 207) = 100.70$, $p < .001$, $\eta^2 =. 33$ and ostracism, $F(1, 207) = 88.32$, $p < .001$, $\eta^2 =. 30$, that was qualified by a significant person x ostracism two-way interaction, $F(1, 207) = 102.63$, $p < .001$, $\eta^2 =.33$. When the sources excluded the target, they were rated as less fair than the target, $F(1, 207) = 202.35$, $p < .001$, $\eta^2 =.50$. When the sources included the target, there was no significant difference in the perception of targets and sources, $F(1, 207) = 0.01$, $p = .946$, $\eta^2 =.00$.

Mediation analyses. We ran two mediation models with MPLUS (Muthén & Muthén, 1998-2015), using 5,000 bootstrap estimates: (a) a mediation model for the effect of conformity on punishment of the target via fairness and evaluation of the *target* and (b) a mediated moderation model for the effect of the ostracism x conformity interaction on punishment of the sources via fairness and evaluation of the *sources*. As for model (a) there was a significant indirect effect of conformity on punishment of the target via evaluation of the target, $b_{indirect} = 1.50$, 90% CI = [1.01; 1.97], but not via fairness, $b_{indirect} = 0.19$, 90% CI =

[-0.31; 0.73]. In contrast, in model (b), the indirect effect of the ostracism x conformity interaction on punishment of the sources via evaluation was not significant, $b_{indirect}$ = -0.19, 90% CI = [-0.41; 0.02]. However, there was an indirect effect of the interaction on punishment via fairness, $b_{indirect}$ = -0.28, 90% CI = [-0.45; -0.11].

Discussion

Study 4 replicated the findings of Study 3 and extended them in at least three important respects. First, the results suggest that the violation of social norms by either the target or sources drives the effects on punishment. This conclusion receives further support by the fact that the effect replicated even though the target behavior could not be perceived as an interpersonal transgression against the sources, thus refuting a potential alternative account of Study 3. In particular, participants perceived the game's injunctive social norm as being cooperative, and agreed with the sources excluding an uncooperative target even if the target had not harmed the sources personally. These findings are largely in line with findings on third-party punishment, which hold that individuals punish uncooperative others even if they are not personally put at a disadvantage (Fehr & Fischbacher, 2004). Second, by incentivizing participants to make punishment decisions in line with the descriptive social norm, we made sure that observers did not merely punish because they could do so without consequences, but that they also took into account how others might react to their decisions. Finally, we explored the role of potential mediators. We found that when the target violated social norms, evaluations seem to be of relevance when making a decision about punishing a norm-violating target. On the other hand, fairness considerations seem to be the more important factor associated with punishing the sources. Although both variables (fairness and evaluation) are highly inter-correlated and the design does not allow for causal conclusions, we offer the cautious speculation that individuals who violate social norms mainly get punished because they are evaluated negatively, while sources are punished when they ostracize a target for no good reason because this act is seen as unfair. Independent of this

speculation, the present findings extend the literature by showing that reactions towards targets versus sources are associated with different concerns (negative evaluations of targets; perceptions of unfairness for sources).

Our suggested theoretical model, which assumes that judgments of fairness as well as evaluations of the target and sources affect participants' sanctioning behavior, is purely theoretically driven. However, as the mediators and the dependent variables were measured at the same time, a different model would also be conceivable, such as sanctioning decisions affecting evaluations. Still, the here proposed mediational chain is consistent with prior research that established a direct link between the ostracism situation and moral judgments (Rudert et al., 2018). In a similar fashion, we also assumed that especially in situations in which participants have no additional information or opinion about the persons they are about to evaluate, it is more likely that attributions and judgments about fairness precede evaluations than vice versa (see also Arpin, Froehlich, Lantian, Rudert, & Stelter, 2017; Chatman & Von Hippel, 2001, for a similar argument).

General Discussion

Observers' reactions to ostracism are not one-sided, but differ depending on the respective social norms: If the target has violated a social norm before, observers sanction the target's behavior. In contrast, if the sources ostracize a target without an apparent socially acceptable reason, observers sanction the sources' behavior by assigning them less money (Studies 1 and 2) or taking money away (Studies 3 and 4). These effects were mediated via negative evaluations and fairness judgments (Studies 3 and 4). Results suggest that evaluation criteria such as participants' anger, sympathy, and willingness to cooperate seem to be more important for the punishment of the target, while fairness considerations seem to be more strongly associated with punishment of the sources (Study 4).

Results suggest that target sanctioning occurred independently of whether the sources ostracized the target or not. Still, in Studies 3 and 4, excluded norm-violating targets were

evaluated particularly negatively and the targets were perceived as highly unfair compared to the sources. The observed difference in evaluation was less extreme in the inclusion group (Study 4). Presumably, observers might perceive inclusion after norm-violation as a signal that the norm violation has been forgiven by the sources, thus resulting in a less negative evaluation of the target. Reversely, exclusion might be taken as a signal that ostracism as a corrective action was necessary because the target had misbehaved. This is in line with theoretical considerations that ostracism serves as a "warning shot" that should make the target adhere to social norms in the future (Kerr et al., 2009). Potentially, whether a person is ostracized by others might serve as an important cue for observers regarding whether they would personally like to interact with that person. The effects were rather small, though, which may be due to the fact that observers knew precisely how the target had acted; research has shown that observers tend to rely on their explicit knowledge about the situation before they consider behavioral cues from others to form a moral judgement (Rudert et al., 2018). Additionally, observers might consider whether sanctioning behavior is socially accepted by others or they are likely to become a target of exclusion themselves if they ostracize others. When in doubt about the social acceptance of their behavior, observers might choose not to sanction even if they judge a behavior as morally wrong. Consistent with this reasoning, we observed that observers' moral judgement was harsher than their actual sanctions in Study 2.

The present studies mostly focus on social norms in moral contexts, such as norms about honesty (Study 1), fidelity (Study 2), as well as cooperation (Studies 3 and 4). We do not investigate the effects of violations of social norms that are merely descriptive norms and do not represent moral transgressions. Our assumption is that especially in situations with a strong inclusion norm and in which violations of the descriptive norms do not hurt the functioning of the group or society in general, it is unlikely that ostracism of such norm-violating targets would be accepted and not sanctioned by observers. In line with this, Rudert and colleagues (2018) showed that ostracizing individuals merely because they are the odd-

one-out in a group is evaluated negatively by observers. It should be noted, however, that norms that are essential for groups or societies often either tend to be moral or become moralized over time (Täuber, 2019), as they represent a central part of a group or society's identity and thus will be protected by the respective members.

General Tendency to Sanction

In Studies 1-3, sanctioning behavior had no personal consequences so participants had neither a direct benefit nor a disadvantage from sanctioning others. An exception is Study 4, in which participants were incentivized to act in line with what others perceived as acceptable. Still, the lack of personal costs may explain why individuals chose to act very generously on average, so that bonuses were relatively high and sanctions relatively mild. Most likely, if participants had needed to distribute the money between themselves and others, the distributions of money to others would have been lower overall. However, given previous research on ostracism and punishment (Will et al., 2013), there is good reason to expect that participants would have behaved very similarly even if they had to invest their own resources to punish in order to restore fairness.

Inspection of sanction tendencies across studies further reveals that sanctions in Studies 1 and 2 were generally lower than in Studies 3 and 4. On a speculative note, we think that this difference arises as a function of the relationship between misbehavior and sanctions. In particular, in Studies 1 and 2, sanctions were not directly related to the misbehavior, and were relatively mild. In Studies 3 and 4, punishment was relatively harsh, presumably because it had a direct relation to the norm-violation (undeserved money is taken away) and participants could restore fairness and an equal distribution of resources with their punishment. Alternatively, the sources and the target were not friends in Studies 3 and 4, so the participants might have felt more entitled to intervene than in a dispute between friends, although that does not explain why they intervened at the expense of the target only. Interestingly, we found no evidence for compensation of norm-consistent targets of ostracism

in any of our studies, even though it has been demonstrated in other research (Wesselmann, Wirth, et al., 2013).

Theoretical and Practical Implications

Observers have received little attention in ostracism research, despite being able to make a critical difference. The judgements as well as the actions or inactions of neutral individuals can legitimize the sources' behavior, or provide support to the target. We speculate that intentional ostracism at school, university, or the work place often involves more individuals than the target and the sources. Understanding observers' considerations and behavior may prove vital for targets and sources, as well as institutions that wish to protect both their culture and individual agents. Most research on observers of ostracism has focused on observers empathizing with the targets and/or punishing the sources for violating the inclusion norm (Güroğlu et al., 2013; Over & Uskul, 2016; Wesselmann, Williams, et al., 2013; Will et al., 2013). In contrast, this contribution demonstrates that observers do not simply condemn and punish ostracism, nor do they mindlessly side with the sources. Instead, observers make distinctive moral judgments in line with the prevailing social norm and specifically sanction violations against the respective social norm.

While this behavior is potentially a strong control mechanism to uphold social norms, it could prove problematic: First, observers often have no information about whether the target has violated a social norm. Observers then revert to heuristics and stereotypes, for instance, based on a target's facial appearance, to make moral judgements (e.g., Rudert, Reutner, et al., 2017). Because some of these cues have little objective validity, these judgments may be prone to error. In the worst case, this could result either in undeserved punishment of innocent targets for norm violations they have not committed, or in observers turning against a group of sources that has merely tried to protect itself from a deviant or selfish target.

Second, our research indicates that observers also take (possible) reactions of others into account. This can be problematic when observers choose not to act in line with their

moral judgement because no one else does. This so-called bystander effect (Chekroun & Brauer, 2002; Latané & Darley, 1969) might be particularly damaging in ostracism situations, because from the target's perspective, it is possibly hard to distinguish between an unresponsive observer and a purposefully ostracizing source.

Finally, in many cases, sanctions might not be the best way to make others adhere to social norms. When the sources feel that they had an acceptable reason for ostracizing, they might react with anger when they are punished. Moreover, targets that feel unfairly ostracized and/or punished will often withdraw or even aggress against the sources (Ren, Wesselmann, & Williams, 2016; Twenge, Baumeister, Tice, & Stucke, 2001).

It should be emphasized that our conclusions are derived from experimental data and may be fruitfully validated in the field. If the findings replicate with strong effect sizes in a more applied context, important implications for persons dealing with ostracism in professional contexts may follow (teachers, counselors, conflict mediators, HR employees). Of key importance might be the ability and motivation to find out the background of an ostracism episode as well as communication about how individuals can deal with norm-violations and ostracism in a constructive way without reverting to sanctions.

Conclusion

Four studies show that observers sanction ostracizing behavior of the sources, but only if they feel that the sources had no socially accepted reason. If the target violated social norms and was ostracized as a consequence, observers sanctioned the target's behavior instead. The results were driven by considerations of fairness as well as different evaluations of targets and sources.

Footnotes

[1] In general, the pattern of the responsibility measure in Studies 1-3 was mostly similar to the findings for punishment but less consistent. As a compromise between exhaustiveness and comprehensibility of the manuscript, we chose to report the findings in an online supplement (https://edoc.unibas.ch/70281/) and do not discuss the measure further.

[2] In all studies, excluding participants from the analyses who failed to correctly answer one or more manipulation checks neither changed the pattern of results nor the significance levels.

[3] In the pre-registration, we stated that we would break down the three-way interaction by target and sources. However, when writing the manuscript, we felt that the direct comparison between target and sources was more important than the direct comparison between the target behavior conditions. Thus, we ultimately decided to split the sample based on the target's behavior in all studies. When splitting the sample by target and sources as suggested in the pre-registration, for the target, there is an effect of target behavior, $F(1, 414) = 310.97$, $p < .001$, $\eta^2 = .43$, while for the sources, there is a 2-way interaction target behavior x ostracism, $F(1, 414) = 18.88$, $p < .001$, $\eta^2 = .04$.

References

Aarts, H., & Dijksterhuis, A. (2003). The silence of the library: Environment, situational norm, and social behavior. *Journal of Personality and Social Psychology, 84*(1), 18 - 28. doi:10.1037/0022-3514.84.1.18

Arpin, S. N., Froehlich, L., Lantian, A., Rudert, S. C., & Stelter, M. (2017). When "We" or "They" exclude others: Attributing and evaluating ostracism observed in in-groups and out-groups. *Comprehensive Results in Social Psychology, 2*(2-3), 143-165. doi:10.1080/23743603.2017.1358477

Blackhart, G. C., Nelson, B. C., Knowles, M. L., & Baumeister, R. F. (2009). Rejection elicits emotional reactions but neither causes immediate distress nor lowers self-esteem: A meta-analytic review of 192 studies on social exclusion. *Personality and Social Psychology Review, 13*(4), 269-309. doi:10.1177/1088868309346065

Chatman, C. M., & Von Hippel, W. (2001). Attributional mediation of in-group bias. *Journal of Experimental Social Psychology, 37*(3), 267-272.

Chekroun, P., & Brauer, M. (2002). The bystander effect and social control behavior: The effect of the presence of others on people's reactions to norm violations. *European Journal of Social Psychology, 32*(6), 853-867. doi:10.1002/ejsp.126

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (4th ed., Vol. 1, 2, pp. 151-192). New York, NY: McGraw-Hill.

Ditrich, L., & Sassenberg, K. (2016). It's either you or me! Impact of deviations on social exclusion and leaving. *Group Processes & Intergroup Relations, 19*(5), 630-652. doi:10.1177/1368430216638533

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2), 175-191. doi:10.3758/BF03193146

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and human behavior, 25*(2), 63-87.

Gibbs, J. P. (1965). Norms: The problem of definition and classification. *American Journal of Sociology, 70*(5), 586-594. doi:10.2307/2774978

Güroğlu, B., Will, G.-J., & Klapwijk, E. T. (2013). Some bullies are more equal than others: Peer relationships modulate altruistic punishment of bullies after observing ostracism. *International Journal of Developmental Science, 7*(1), 13-23. doi:10.3233/DEV-1312117

Hartgerink, C. H., van Beest, I., Wicherts, J. M., & Williams, K. D. (2015). The ordinal effects of ostracism: A meta-analysis of 120 Cyberball studies. *PloS one, 10*(5), e0127002. doi:10.1371/journal.pone.0127002

Hayes, A. F. (2013). *Introduction to mediation, moderation and conditional process analysis.* New York, NY: Guilford.

Kerr, N. L., Rumble, A. C., Park, E. S., Ouwerkerk, J. W., Parks, C. D., Gallucci, M., & van Lange, P. A. M. (2009). "How many bad apples does it take to spoil the whole barrel?": Social exclusion and toleration for bad apples. *Journal of Experimental Social Psychology, 45*(4), 603-613. doi:10.1016/j.jesp.2009.02.017

Latané, B., & Darley, J. M. (1969). Bystander "Apathy". *American Scientist, 57*(2), 244-268.

Masten, C. L., Morelli, S. A., & Eisenberger, N. I. (2011). An fMRI investigation of empathy for 'social pain' and subsequent prosocial behavior. *NeuroImage, 55*(1), 381-388. doi:10.1016/j.neuroimage.2010.11.060

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus User's Guide* (7 ed.). Los Angeles, CA: Muthén & Muthén.

Over, H., & Uskul, A. K. (2016). Culture moderates children's responses to ostracism situations. *Journal of Personality and Social Psychology, 110*(5), 710 - 724. doi:10.1037/pspi0000050

Pryor, J. B., Reeder, G. D., Wesselmann, E. D., Williams, K. D., & Wirth, J. H. (2013). The

    influence of social norms upon behavioral expressions of implicit and explicit weight-

    related stigma in an interactive game. *The Yale journal of biology and medicine, 86*(2),

    189 - 201.

Ren, D., Wesselmann, E., & Williams, K. D. (2016). Evidence for another response to

    ostracism: Solitude seeking. *Social Psychological and Personality Science, 7*(3), 204-

    212. doi:10.1177/1948550615616169

Rudert, S. C., & Greifeneder, R. (2016). When it's okay that I don't play: Social norms and the

    situated construal of social exclusion. *Personality and Social Psychology Bulletin,

    42*(7), 955-969. doi:10.1177/0146167216649606

Rudert, S. C., & Greifeneder, R. (2019). Observing Ostracism: How observers interpret and

    respond to ostracism situations In S. C. Rudert, R. Greifeneder, & K. D. Williams

    (Eds.), *Current Directions in Ostracism, Social Exclusion and Rejection* (pp. 136-

    154). Oxford, UK: Routledge.

Rudert, S. C., Hales, A. H., Greifeneder, R., & Williams, K. D. (2017). When silence is not

    golden: Why acknowledgement matters even when being excluded. *Personality and

    Social Psychology Bulletin, 43*(5), 678 - 692. doi:10.1177/0146167217695554

Rudert, S. C., Reutner, L., Greifeneder, R., & Walker, M. (2017). Faced with exclusion:

    Perceived facial warmth and competence influence moral judgments of social

    exclusion. *Journal of Experimental Social Psychology, 68*, 101-112.

    doi:10.1016/j.jesp.2016.06.005

Rudert, S. C., Sutter, D., Corrodi, C., & Greifeneder, R. (2018). Who's to blame?

    Dissimilarity as a cue in moral judgments of observed ostracism episodes. *Journal of

    Personality and Social Psychology, 115*, 31-53. doi:10.1037/pspa0000122

Skarlicki, D. P., O'Reilly, J., & Kulik, C. T. (2015). The third-party perspective of (in) justice. In R. Cropanzano & M. L. Ambrose (Eds.), *The Oxford Handbook of Justice in the Workplace* (pp. 235 - 255). Oxford: Oxford University Press.

Spada, H., & Opwis, K. (1985). Ökologisches Handeln im Konflikt: Die Allmende-Klemme. In P. Day, U. Fuhrer, & U. Laucken (Eds.), *Umwelt und Handeln* (pp. 63-85). Tübingen: Attempto.

Stamkou, E., van Kleef, G. A., Homan, A. C., & Galinsky, A. D. (2016). How norm violations shape social hierarchies: Those who stand on top block norm violators from rising up. *Group Processes & Intergroup Relations, 19*(5), 608-629. doi:10.1177/1368430216641305

Täuber, S. (2019). Moralization as legitimization for ostracism: Effects on intergroup dynamics and social cohesion. In S. C. Rudert, R. Greifeneder, & K. D. Williams (Eds.), *Current Directions in Ostracism, Social Exclusion and Rejection.* Abingdon, UK: Routledge.

Tuscherer, T., Sacco, D. F., Wirth, J. H., Claypool, H. M., Hugenberg, K., & Wesselmann, E. D. (2015). Responses to exclusion are moderated by its perceived fairness. *European Journal of Social Psychology, 46*, 280-293. doi:10.1002/ejsp.2152

Twenge, J. M., Baumeister, R. F., Tice, D. M., & Stucke, T. S. (2001). If you can't join them, beat them: Effects of social exclusion on aggressive behavior. *Journal of Personality and Social Psychology, 81*(6), 1058-1069. doi:10.1037/0022-3514.81.6.1058

Wesselmann, E. D., Bagg, D., & Williams, K. D. (2009). "I feel your pain": The effects of observing ostracism on the ostracism detection system. *Journal of Experimental Social Psychology, 45*(6), 1308-1311. doi:10.1016/j.jesp.2009.08.003

Wesselmann, E. D., Butler, F. A., Williams, K. D., & Pickett, C. L. (2010). Adding injury to insult: Unexpected rejection leads to more aggressive responses. *Aggressive Behavior, 36*(4), 232-237. doi:10.1002/ab.20347

Wesselmann, E. D., Williams, K. D., & Hales, A. H. (2013). Vicarious ostracism. *Frontiers in Human Neuroscience, 7*. doi:10.3389/fnhum.2013.00153

Wesselmann, E. D., Wirth, J. H., & Bernstein, M. J. (2017). Expectations of Social Inclusion and Exclusion. *Frontiers in Psychology, 8*(112). doi:10.3389/fpsyg.2017.00112

Wesselmann, E. D., Wirth, J. H., Pryor, J. B., Reeder, G. D., & Williams, K. D. (2013). When do we ostracize? *Social Psychological and Personality Science, 4*(1), 108-115. doi:10.1177/1948550612443386

Wesselmann, E. D., Wirth, J. H., Pryor, J. B., Reeder, G. D., & Williams, K. D. (2015). The role of burden and deviation in ostracizing others. *The Journal of Social Psychology, 155*(5), 483-496.

Will, G.-J., Crone, E. A., van den Bos, W., & Güroğlu, B. (2013). Acting on observed social exclusion: Developmental perspectives on punishment of excluders and compensation of victims. *Developmental Psychology, 49*(12), 2236-2244. doi:10.1037/a0032299

Williams, K. D. (2009). Ostracism: A temporal need-threat model. In P. Z. Mark (Ed.), *Advances in experimental social psychology* (Vol. 41, pp. 275-314). San Diego, CA: Elsevier Academic Press.

Williams, K. D., Cheung, C. K., & Choi, W. (2000). Cyberostracism: Effects of being ignored over the Internet. *Journal of Personality and Social Psychology, 79*(5), 748-762. doi:10.1037/0022-3514.79.5.748

Williams, K. D., Shore, W. J., & Grahe, J. E. (1998). The silent treatment: Perceptions of its behaviors and associated feelings. *Group Processes & Intergroup Relations, 1*(2), 117-141. doi:10.1177/1368430298012002

Table 1

Results for Study 1.

Dependent Variable	Repeated Measure	Norm-violating Target		Norm-consistent Target	
		Exclusion	*Inclusion*	*Exclusion*	*Inclusion*
Money (CHF)	Target	3.07 [c] (1.66)	3.41 [bc] (1.45)	3.82 [b] (1.45)	4.00 [b] (1.45)
	Sources	3.78 [ab] (1.47)	3.78 [b] (1.29)	3.19 [a] (1.47)	3.97 [b] (1.44)
Participation	Target	2.20 [a] (1.42)	4.67 [b] (1.88)	2.06 [a] (1.23)	5.13 [b] (1.36)
	Sources	6.28 [c] (.99)	5.33 [b] (.96)	6.57 [c] (.75)	5.18 [b] (1.00)
Excluded - Included	Target	2.48 [a] (1.15)	4.85 [b] (1.49)	2.29 [a] (1.00)	5.43 [b] (1.29)
	Sources	6.31 [c] (1.31)	5.28 [b] (1.00)	6.22 [c] (1.29)	5.23 [b] (1.45)
Ignored - Acknowledged	Target	2.40 [a] (1.35)	4.44 [c] (2.03)	2.38 [a] (1.37)	5.18 [c] (1.70)
	Sources	6.38 [b] (1.10)	5.19 [c] (1.33)	6.22 [b] (1.19)	5.41 [c] (1.26)
Lie Acceptable	Target	2.57 [b] (1.48)	3.07 [a] (.85)	3.53 [a] (2.02)	3.97 [a] (1.77)
Lie Desirable	Target	2.80 [a] (1.13)	2.62 [a] (.70)	2.41 [a] (1.35)	2.53 [a] (1.25)
Lie Morally Good	Target	2.33 [b] (1.49)	2.62 [b] (1.13)	3.62 [a] (1.71)	3.73 [a] (1.55)

Note. Means (and standard deviations) as a function of the four experimental conditions, separately for targets and sources of ostracism. The letters a - d represent significant differences between groups; all values in the same column or row that share the same letter do not differ significantly from each other, values with different letters do.

Table 2

Results for Study 2.

Dependent Variable	Repeated Measure	Norm-violating Target (strong)		Norm-violating Target (weak)		Norm-consistent Target	
		Exclusion	*Control*	*Exclusion*	*Control*	*Exclusion*	*Control*
Money (CHF)	Target	3.10 [ac](1.82)	2.34 [c] (1.90)	3.70 [ab] (1.40)	3.97 [bd] (1.40)	3.45 [ab] (1.76)	4.00 [b] (1.44)
	Sources	3.74 [b] (1.45)	3.41 [ab] (1.40)	3.33 [ab] (1.61)	4.42 [d] (1.02)	2.23 [c] (1.49)	4.03 [abd] (1.21)
Participation	Target	2.21 [a] (.77)		2.50 [a] (1.50)		1.91 [a] (.61)	
	Sources	6.72 [b] (.64)		6.63 [b] (.68)		6.84 [b] (.32)	
Excluded - Included	Target	2.43 [a] (.50)		2.83 [a] (1.24)		2.36 [a] (.73)	
	Sources	6.82 [b] (.35)		6.56 [b] (.84)		6.57 [b] (.62)	
Ignored - Acknowledged	Target	2.65 [a] (.89)		2.96 [a] (1.30)		2.67 [a] (.91)	
	Sources	6.75 [b] (.51)		6.56 [b] (.83)		6.55 [b] (.61)	
Morally Good	Target	2.60 [a] (.83)	2.94 [a] (1.24)	3.08 [a] (1.42)	2.82 [a] (1.29)	5.00 [b] (1.84)	5.12 [b] (1.78)

Note. Means (and standard deviations) as a function of the six experimental conditions, separately for targets and sources of ostracism. The letters a - d represent significant differences between groups; all values in the same column or row that share the same letter do not differ significantly from each other, values with different letters do.

Table 3

Results for Study 3.

Dependent Variable	Repeated Measure	Norm-violating Target		Norm-consistent Target	
		Exclusion	*Inclusion*	*Exclusion*	*Inclusion*
Subtracted Money (£)	Target	1.29 [a] (.89)	1.09 [a] (.92)	.15 [b] (.45)	.21 [b] (.51)
	Sources	.08 [b] (.40)	.13 [b] (.43)	.37 [a] (.56)	.20 [ab] (.49)
Evaluation	Target	2.93 [c] (1.64)	3.57 [b] (1.54)	6.01 [a] (0.96)	5.91 [a] (1.11)
	Sources	6.21 [a] (1.04)	5.94 [a] (1.07)	4.02 [b] (1.52)	5.89 [a] (1.06)

Note. Means (and standard deviations) as a function of the four experimental conditions, separately for targets and sources of ostracism. The letters a - d represent significant differences between groups; all values in the same column or row that share the same letter do not differ significantly from each other, values with different letters do.

Table 4

Results for Study 4.

Dependent Variable	Repeated Measure	Norm-violating Target		Norm-consistent Target	
		Exclusion	*Inclusion*	*Exclusion*	*Inclusion*
Subtracted Money (p)	Target	6.67 [a] (4.09)	6.22 [a] (3.93)	0.83 [b] (2.16)	0.90 [b] (2.26)
	Sources	0.61 [b] (1.96)	0.95 [b] (2.26)	2.40 [c] (2.89)	0.78 [ab] (2.01)
Evaluation	Target	2.71 [a] (1.39)	2.95 [a] (1.52)	6.08 [b] (1.06)	6.13 [b] (.79)
	Sources	6.10 [b] (.91)	5.51 [d] (1.05)	3.99 [c] (1.43)	6.04 [b] (.86)
Fairness	Target	2.14 [a] (1.56)	2.70 [c] (1.67)	6.36 [b] (1.18)	6.61 [b] (.98)
	Sources	6.10 [b] (1.11)	5.81 [b] (1.24)	4.34 [d] (1.67)	6.62 [b] (.88)

Note. Means (and standard deviations) as a function of the four experimental conditions, separately for targets and sources of ostracism. The letters a - d represent significant differences between groups; all values in the same column or row that share the same letter do not differ significantly from each other, values with different letters do.
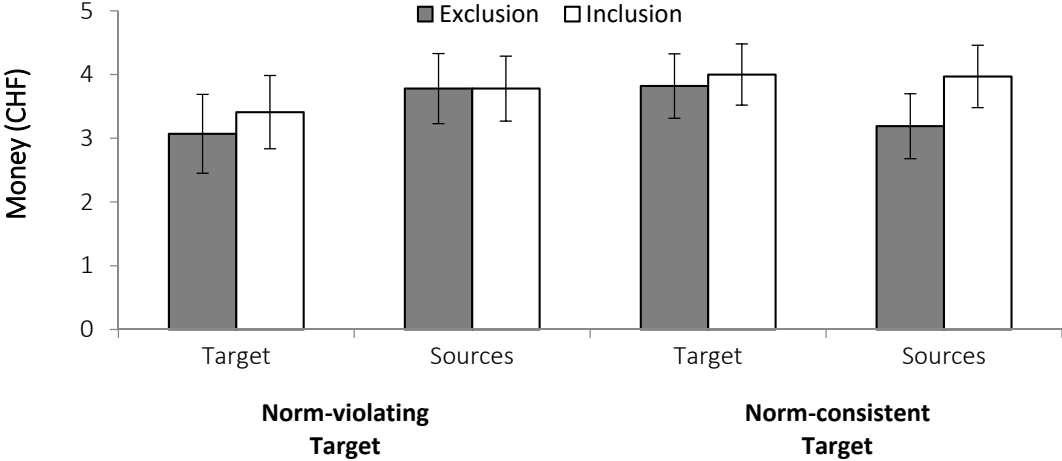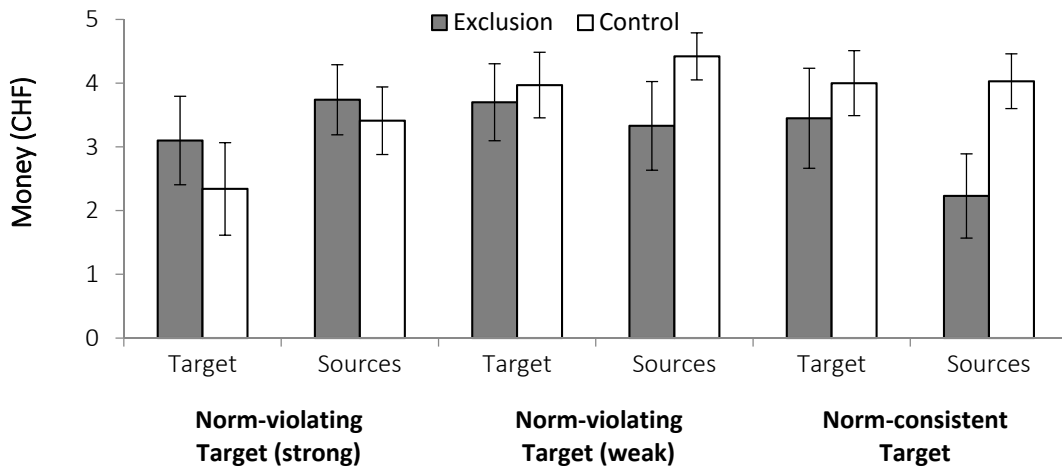
Figure 1. Mean levels of the distributed amount of money for targets and sources (with 95%

Confidence Intervals) as a function of ostracism and norm consistency in Study 1. The

exclusion condition is displayed as grey bars and the inclusion condition as white bars.

Figure 2. Mean levels of the distributed amount of money for targets and sources (with 95%

Confidence Intervals) as a function of ostracism and norm consistency in Study 2. The

exclusion condition is displayed as grey bars and the control condition as white bars.
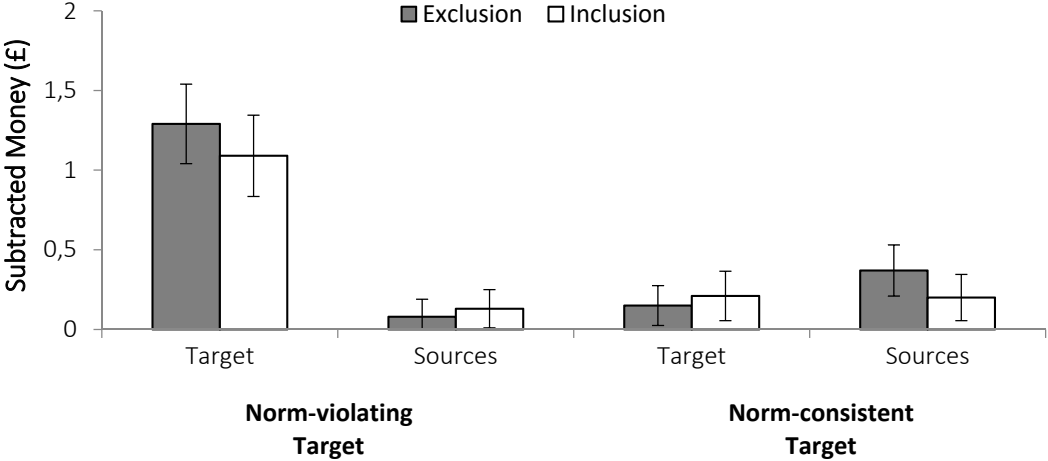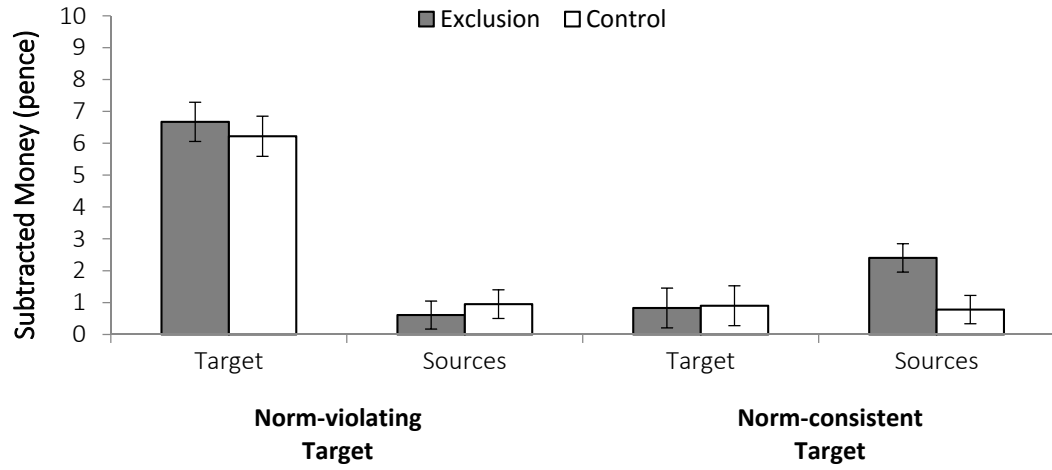
Figure 3. Mean levels of the amount of money subtracted from targets and sources (with 95%

Confidence Intervals) as a function of ostracism and norm consistency in Study 3. The

exclusion condition is displayed as grey bars and the inclusion condition as white bars.

Figure 4. Mean levels of the amount of money subtracted from targets and sources (with 95%

Confidence Intervals) as a function of ostracism and norm consistency in Study 4. The

exclusion condition is displayed as grey bars and the inclusion condition as white bars.