Running head: OBSERVED OSTRACISM

Who's to Blame?

Dissimilarity as a Cue in Moral Judgments of Observed Ostracism Episodes

Selma Carolin Rudert and Daniela Sutter

University of Basel, Switzerland

Veronique Charlotte Corrodi

Pädagogische Hochschule St. Gallen, Switzerland

Rainer Greifeneder

University of Basel, Switzerland

Author Note

Correspondence concerning this article should be addressed to Selma C. Rudert, Department of Social Psychology, University of Basel, Missionsstrasse 64 a, CH-4055 Basel, Switzerland, E-Mail: selma.rudert@unibas.ch

Abstract

When observing an ostracism episode, observers may wish to know whether ostracism is justified or not. If ostracism appears unjustified, observers will likely blame the sources and sympathize with the target; if it appears justified, observers will likely blame and devalue the target. Here we introduce the "social dissimilarity rule," which holds that observers base their moral judgments on dissimilarities between the members of the observed group. In five studies, participants either recalled observed ostracism episodes or observed group interactions in which one group member was ostracized (e.g., in a chat or a group-working task). Results show that if similar persons exclude a dissimilar target (target is an "odd-one-out"), observers attribute ostracism to malicious motives of the ostracizers, such as ingroup favoritism, and devalue the ostracizers. However, if ostracism cannot be explained by social dissimilarity between the sources and the target, observers assume that the target is being punished for a norm deviation (punitive motive) and devalue the target. Use of the social dissimilarity rule was neither moderated by cognitive load (Study 3) nor by the perceived essentiality of the group distinction (Study 4). But if participants knew that the target previously deviated from a norm, knowledge about the situation had a stronger effect on moral judgments (Study 5) than social dissimilarity. These findings further our understanding of how observers make moral judgments about ostracism, which is important given that an observer's moral judgment can strongly impact bystander behavior and thus target recovery and well-being.

Keywords: Ostracism, Social Exclusion, Observation, Attribution, Punishment

Who's to Blame?

Dissimilarity as a Cue in Moral Judgments of Observed Ostracism Episodes

No matter whether you think of, for instance, a child being ignored and excluded by his or her peers, a colleague from work not being asked to join for lunch, a homeless beggar deliberately being overlooked in the street, or maybe even a family member not being invited to gatherings because s/he always causes disagreements - social exclusion and ostracism can be observed everyday and in almost every social context. However, individuals' reactions when observing social ostracism differ vastly. Sometimes, observers help and support the ostracized target, such as by inviting a colleague who has to sit alone during lunch to their own table. At other times, observers may choose not to get involved in the situation or even exclude the ostracized individual themselves. Taking into account the detrimental effects of ostracism on a variety of motivational, cognitive, affective, physiological, and behavioral variables (e.g., Williams, 2009), as well as the strong positive impact that social support and acknowledgement have on the well-being of ostracized targets (e.g., Eck, Schoel, & Greifeneder, 2016; Rudert, Hales, Greifeneder, & Williams, 2017; Rudert, Janke, & Greifeneder, 2017; Teng & Chen, 2012; Zhou & Gao, 2008), understanding when and why observers decide to support ostracized others is of high importance.

The present contribution focuses on how *observers* of ostracism interpret and react to situations of social ostracism. We start from the assumption that individuals who observe an ostracism episode make an attribution about the underlying motives of the ostracizing group and then form a moral judgment about who is to blame. Two outcomes are particularly conceivable: on the one hand, ostracism can be attributed to a malicious motive. In this case, ostracism is likely judged as an inappropriate, unacceptable behavior of the ostracizing group, and observers therefore

derogate the ostracizing group and empathize with the target. On the other hand, ostracism can be attributed to a punitive motive, for instance, because the target's behavior violates standard social norms. In this case, ostracism is likely judged as acceptable, and observers empathize with the ostracizing group, and derogate the target.

Which of the two pathways is more likely is governed, for instance, by the observer's knowledge of the situation or the characteristics of those involved in the ostracism episode. Here we focus on the latter, namely characteristics of those who ostracize (the sources) and characteristics of those who are ostracized (the targets). We argue that similarities or differences between apparent characteristics constitute a key social cue that individuals rely on in the absence of more diagnostic information. In what we refer to as the *social dissimilarity rule*, we state that when the target is dissimilar to the sources, observers tend to attribute ostracism to a malicious motive, derogate the sources, and empathize with the target. In contrast, when the social dissimilarity rule is not applicable (e.g., because the target is similar to the sources), observers tend to attribute ostracism to a punitive motive, derogate the target, and empathize with the sources. In what follows, we provide the theoretical background for this reasoning and detail the suggested hypotheses.

Observing a Social Exclusion Episode: Malicious and Punitive Motives

The first chain element in our argument holds that observers tend to perceive ostracism as being due to either a malicious or punitive motive. A *malicious attribution* means that ostracism is perceived as an inappropriate, unacceptable behavior by the ostracizing group. This attribution is likely to result from a strong shared default to include others, which prevails in many situations (Rudert & Greifeneder, 2016; Wesselmann, Wirth, Pryor, Reeder, & Williams, 2013). As a result of this default, observers will attribute ostracism to malicious motives of the sources, disapprove

of their behavior and side with the target. In line with this assumption, research has shown that observing ostracism in the virtual ball-throwing game Cyberball makes individuals feel uncomfortable and threatened as well as makes them empathize with the targets (Masten, Morelli, & Eisenberger, 2011; Wesselmann, Bagg, & Williams, 2009; Wesselmann, Williams, & Hales, 2013; Will, Crone, van den Bos, & Güroğlu, 2013). Given the chance, observers even attempt to compensate the target, for instance, by directing more throws towards him/her (Wesselmann, Wirth, et al., 2013).

However, other studies show that there are situations in which individuals attribute ostracism to a *punitive motive* of the sources, too. We define a punitive motive as being void of selfish or malicious desire, but instead reflecting the intent to either correct the target's behavior or protect the group. If, for instance, the observers know that the ostracized target is a troublemaker, then excluding him or her may be perceived as a fair punishment. In this case, ostracism represents an apparently justified sanction of the target's previous norm deviation with the aim to restore harmony in the group, strengthen the group norm, and discourage the troublemaker and/or other members from deviating from the norm as well (Ditrich & Sassenberg, 2016; Kerr et al., 2009). Wesselmann and colleagues (2013) demonstrated empirical evidence for punitive ostracism, showing that observers' support for the excluded target in Cyberball decreased when the target's throws were slower than the other players' throws. As soon as participants felt that the target was being a burden, they stopped compensating him/her and even joined the sources in ostracizing the target. In a follow-up study, the authors demonstrated that participants' attributions for the slowness of an ostracized Cyberball-player varied, ranging from external attributions such as low speed of the target's internet connection to internal attributions such as laziness of the target (Wesselmann, Williams, & Wirth, 2014). These findings are also in line with

Weiner's attribution theory (Weiner, 2006), stating that especially when individuals attribute the occurrence of a negative event to the target's behavior (e.g., the target did not invest enough effort), the target is perceived as responsible for the event (see also Arpin, Froehlich, Lantian, Rudert, & Stelter, 2017).

Attributing Motives When There is Little Diagnostic Information

In previous studies that investigated possible reasons for ostracism, participants usually knew whether an individual has transgressed against a group norm before, because they were part of the group themselves and witnessed the target deviating from the norm (Ditrich & Sassenberg, 2016; Gooley, Zadro, Williams, Svetieva, & Gonsalkorale, 2015; Wesselmann, Wirth, et al., 2013). However, in real life, observers of ostracism often lack relevant information about the situation or the target, especially when the observer is not even a part of the respective group. Observers may also not be able to fully process an ostracism situation due to limited motivation as well as cognitive capacity (Petty & Cacioppo, 1984). To illustrate, think of a teacher who is monitoring the schoolyard during a recess and observes a group of unknown children ostracizing another unknown child. Given that she has no information about the group or the previous history, how would she make a judgment about whether she should intervene or not?

Hence, the second chain element in our argument entails that individuals making a moral judgment about observed ostracism episodes will base this judgment on the presence of apparent reasons for excluding others, which may or may not justify ostracism in the eye of the observer. In situations in which valid episodic information is sparse or missing, observers likely make use of heuristics and social stereotypes (Brewer, 1988; Fiske & Neuberg, 1990; Macrae, Milne, & Bodenhausen, 1994). Initial evidence for this was provided by Rudert, Reutner, Greifeneder, and Walker (2017), who have demonstrated that the facial appearance of target can influence

observer's moral judgments, such that ostracizing cold and incompetent looking individuals is perceived as more acceptable. In the present research, we propose that social dissimilarity between the members of the observed group may serve as another heuristic.

Observed Situation: The Target is Ostracized by Dissimilar Others

We assume that observers will attribute ostracism to a malicious motive, namely discrimination, when they see a dissimilar person being ostracized by two similar others. This is because individuals have stereotypical conceptions of prejudice and discrimination (Inman & Baron, 1996) as well as lay theories and expectations with regard to intergroup conflicts and biases (Sommers & Norton, 2006; Vivian & Berkowitz, 1992) from an early age (McKown & Weinstein, 2003). Thus, if observers notice that a rather homogeneous group ostracizes a dissimilar person, one obvious attribution is that the sources are showing ingroup favoritism, and disadvantage an outgroup member. As a consequence of this attribution, we propose that observers will evaluate ostracism as unfair. This is because individuals generally try to avoid being seen as prejudiced (Moss-Racusin, Phelan, & Rudman, 2010), and tend to perceive openly prejudiced persons such as racists in a highly negative way (Crandall, Eshleman, & O'Brien, 2002; Sommers & Norton, 2006). Thus, if the sources are similar and the target is dissimilar from the sources, we assume that observers tend to assume a malicious motive, devalue the sources, and express sympathy for the ostracized target, because it is wrong to ostracize someone just because she/he is the odd-one-out.

Observed Situation: The Target is Ostracized by Similar Others

If the target is ostracized by one or more similar others, the social dissimilarity rule is inapplicable, and thus, observers need an alternative explanation. We argue that, given a sufficiently ambiguous situation, observers will assume that the target is excluded because s/he

committed some norm transgression and is now being punished for it. This attribution is in line with research on the Black Sheep Effect, which states that ingroup members strongly derogate and reject members from their own group who deviate from norms (Marques & Paez, 1994; Marques, Yzerbyt, & Leyens, 1988). Such an attribution may further be self-serving for observers. This is because in highly ambiguous situations, it may be safer to be on the side of the (excluding) majority than risk siding with a potentially deviant target. Thus, we predict that if the social dissimilarity rule cannot be applied to an exclusion situation, observers will tend to blame and devalue the target for his/her current situation.

Overview of the Present Research

In the present contribution, we test the influence of (dis)similarity on observers' attributions and evaluations in ambiguous ostracism situations. In particular, we measure two different attributions: malicious and punitive attributions. Evaluation of target and sources is assessed as the amount of blame, anger, sympathy and interest in cooperating with the respective evaluated person. We assume that observers draw inferences based on the social dissimilarity rule: if the ostracized target is an "odd-one-out," that is, dissimilar to the ostracizing sources, observers tend to attribute ostracism to ingroup favoritism and malicious motives of the sources. As a consequence, observers will sympathize with the target and devalue the sources. In contrast, if the target does not seem to be an "odd-one-out," the social dissimilarity rule cannot be applied. As a result, observers tend to assume that the target violated a social norm, attribute ostracism to a punitive motive, and consequently blame and devalue the target instead.

We test these conjectures in five studies. In Studies 1a and 1b, we investigate which attributions individuals spontaneously make for observed ostracism and thus asked participants to describe ostracism situations that they had observed in real life. These reasons were then

coded by independent raters (Study 1a) and four of them were evaluated by another group of

participants (Study 1b). In the subsequent studies, we manipulated similarity, using vignettes

(Studies 2 and 3) as well as a newly developed experimental chatroom paradigm (Studies 4 and

5) to create an ostracism situation. Similarity between the agents was manipulated via support of

fictitious soccer teams (Studies 2 and 3), ethnic background and hairstyle (Study 4), and

nationality (Study 5). Studies 2 - 4 test whether the effect of similarity on evaluation is mediated

via attributions to a malicious or a punitive motive. Studies 3 and 4 additionally test whether

cognitive load and perceived essentiality of the group affects observers' use of the social

dissimilarity rule. Finally, to test the boundary conditions of the social dissimilarity rule, in

Study 5 we investigate whether observers still apply the social dissimilarity rule when they have

a more informative cue, namely knowledge about the history of the situation.

Study 1a

An important premise of our theoretical assumptions is that individuals perceive

dissimilarity of the ostracized target as a plausible explanation for ostracism in certain situations.

Thus, in a first, preliminary study we assessed which attributions individuals spontaneously

make for observed ostracism, or more specifically, how frequently individuals attribute observed

ostracism episodes to the ostracized target being an "odd-one-out." To do so, we recruited 30

US-American participants (15 females, $M_{age} = 33.34$, $SD = 11.40$) from Prolific Academic for a

payment of £0.50 and asked them to describe a situation in which they observed another person

being ignored and excluded. We further asked participants to state the reasons why they thought

this specific person was ignored and excluded. Below are two examples of answers that

participants provided:

"The last time I can remember someone being purposely ignored and excluded was in elementary school. I went to a small parochial school. There was a boy in my grade who had some cognitive issues and since it was such a small school, his differences were amplified. Looking back, he likely had severe ADD along with some other issues affecting behavior. He rarely was invited to birthday parties and often sat alone at lunch. He was very different from the rest of the students, and since we were all in such a bubble, most people were cruel to him."

"In highschool there was a girl that had wanted to join our group project but the group did not want her to be. Our group was most familiar to her as she used to be welcome. She did not know many others in the class and had to group with unfamiliar people. Our group ignore [sic] her and did not let her join. This girl had previously been part of our group, however she burned that bridge when she talked behind another person's back and shared some secrets. No one trusted her after."

Answers were coded by two independent raters, who were blind to the research question, with respect to the following categories: "target was different," "target behaved inappropriately," "other reason," (Cohen's Kappa = .75, a "substantial agreement" according to Landis & Koch, 1977). The reason for ostracism that was most frequently chosen by both raters was that the ostracized target was dissimilar from the others (Rater 1: 67%; Rater 2: 57%), followed by the ostracized target behaving inappropriately (Rater 1: 27%; Rater 2: 30%). Thus, it seems that individuals attribute ostracism frequently to the fact that the ostracized person is dissimilar to the rest of the group. However, they also seem to be aware that ostracism can be due to punishment for norm violations.

Study 1b

We further assumed that individuals would judge it as wrong to ostracize an individual just because s/he is dissimilar, but would perceive ostracism as more acceptable when the target had previously committed a norm violation. More specifically, ostracism due to a norm violation should be perceived as more justified than ostracism due to dissimilarities. Moreover, individuals who observe a norm-violating target being ostracized should be more angry and less sympathetic towards the target as well as less angry and less sad about the situation compared to when the target has not violated any norms. To test this assumption, we selected four of the situations that individuals had described and presented them to a new group of participants.

Methods

We recruited 50 participants (25 females, $M_{age} = 34.28$, $SD = 12.52$) from Prolific Academic for a payment of £0.50. Sample size was calculated to detect medium-sized main effects of the two different reasons for ostracism on a single dependent variable ($d = .50$, power $= .90$, required $n = 44$). All participants read two of the situations from Study 1a that had been coded by both raters as "target violates a social norm" (constituting a reason for a punitive motive), and two situations that had been coded by both raters as "target is the odd-one-out." (constituting a reason for a malicious motive). The four scenarios were chosen because they seemed well-suited to represent ostracism due to a norm violation versus dissimilarity, plus they all had in common that they referred to ostracism among children or adolescents. After reading a scenario, participants rated the respective scenario on the following five items *"I think it was justified that the person was ignored and excluded," "I feel sympathy with the ignored and excluded person," "I feel angry with the ignored and excluded person," "The situation makes me feel angry,"* and *"The situation makes me feel sad"* (1 = not at all; 7 = very much).

Results

A MANOVA on all five evaluation items with all four situations as repeated measures revealed that the situations differed significantly in evaluation, $F(3, 47) = 19.24$, $p < .001$, $\eta^2 = .39$. We specified contrasts, testing the two situations that were coded as due to the target committing a norm violation against the two situations that were coded as due to the target being dissimilar from the group (1 1 -1 -1). In line with our assumptions, reading a scenario in which the target was excluded because of a norm violation resulted in participants evaluating ostracism as more justified, $F(1, 49) = 253.17$, $p < .001$, $\eta^2 = .84$, compared to reading a scenario in which the target was excluded because s/he was dissimilar from the group. Regarding the evaluation of the target, participants felt more anger, $F(1, 49) = 8.06$, $p = .007$, $\eta^2 = .14$, and less sympathy, $F(1, 49) = 120.42$, $p < .001$, $\eta^2 = .56$, with the excluded and ignored person when reading the norm-violating compared to the dissimilarity scenarios. Regarding the evaluation of the situation, ostracism due to a norm violation made participants feel less angry, $F(1, 49) = 110.31$, $p < .001$, $\eta^2 = .70$, and sad, $F(1, 49) = 124.99$, $p < .001$, $\eta^2 = .72$ than ostracism due to the target being dissimilar from the group. For the descriptive statistics, see *Table 1*.

Discussion

Studies 1a and 1b provide first evidence for our assumption that individuals commonly conceive ostracism situations in terms of two prominent attributions: (1) ostracism is perceived to result from the target being an odd-one-out (malicious ostracism); (2) ostracism is perceived as a punishment of the target for a norm violation (punitive ostracism). The majority of ostracism situations that participants recalled from their real lives could be assigned to one of the two reasons. In the follow-up Study 1b, another group of participants evaluated ostracism due to a norm violation as more justified and reported that they felt more angry and less sympathetic with

the target. Moreover, they were less angry and sad about these situations than about the situations in which a target was ostracized due to being the odd-one-out.

The two studies indicate that observers distinguish between punitive and malicious attributions for ostracism in real life. Also, observers seem to base their moral judgment on these attributions, insofar that ostracism situations that are attributed to a punitive motive are evaluated as more acceptable than ostracism situations that are attributed to a malicious motive. These findings lay the foundation for the subsequent studies, in which we present participants with an ostracism situation and manipulate dissimilarity in an experimental setting.

Study 2

Study 2 was a vignette study, which aimed to find evidence for the social dissimilarity rule. We presented participants with different vignettes, in which either the target, or one of the sources, or no one was the odd-one-out. In line with our assumptions, we expected that when the target was the odd-one-out, participants would attribute ostracism to a malicious motive of the sources and devalue the sources as a consequence. When the target was not the odd-one-out, we expected participants to attribute ostracism to a punitive motive of the sources and devalue the target as a consequence.

Method

Participants and design. Participants were recruited online from Amazon's Mechanical Turk (US Americans only, HIT Approval Rate greater than 96) for a payment of $0.75. We calculated the sample size such as to detect a medium-sized similarity x evaluated person interaction effect on a single dependent variable ($f = .25$, power $= .80$, required $n = 120$). One hundred and fifty-two participants finished the questionnaire (62 females, $M_{age} = 35.59$, $SD = 12.51$). Participants were randomly assigned to three similarity conditions: the target as the odd-

one-out, one of the sources as the odd-one-out, or no odd-one-out. Because the target is the odd-one-out in the first condition only, we expected to find attribution to a punitive motive and devaluation of the target in both the sources as the odd-one-out and the no odd-one-out conditions.

Materials and procedure. Participants were told that they would be presented with a specific part of an interaction sequence between three football fans and that their task would be to evaluate the interaction between the three persons. Next, participants were presented with three football fans called Michael (the target), Peter, and Tom (the sources), and the teams these fans ostensibly supported. To manipulate social similarity, one third of the participants were told that Peter and Tom supported Team Blue while Michael supported Team Red, so that the sources were similar and the target was the odd-one-out. Another third was told that both Michael and Tom supported Team Blue, while Peter supported Team Red, so that the target and one of the sources were similar and the other source was the odd-one-out. The remaining third was told that all three football fans supported Team Blue (no odd-one-out).

Participants were further told that the three football fans had played a cooperation game first. Participants were not informed about the game's outcome, though, so that the social situation was highly ambiguous: Although observing participants knew that the group had interacted before, they had no idea about how this interaction went. Ostensibly, right after the cooperation game, each of the three football fans had had the option to decide whether he wanted to work together with the others in a subsequent group task or not. Participants were then presented with the ostensible choices of the three persons: While the target stated that he had no preferences, the sources both declared that they did not want to work together with the target.

Participants evaluated all three football fans on the dependent variables. Blame was assessed by asking: "*Please indicate for each person how likely you think it is that he set the events in motion that eventually resulted in the choices about the group task.*" (1 = not likely at all, 7 = very likely). Moreover, participants were asked: "*Assuming you had to play the cooperation game yourself: How much would you like to play with each of the three persons?*" as well as "*I feel angry about [name]'s behavior*" and " *I can sympathize with [name]*" (1 = not at all, 7 = very much; items for anger and sympathy derived from Wesselmann, Wirth, et al., 2013). The evaluations for the sources were averaged separately for each variable (Spearman's ρ = .95 - .97).

Additionally, we assessed attributions that participants made about the observed exclusion situation. We presented participants with six statements, each starting with "I think that…" Attribution to a punitive motive was assessed with the following three items: "*…Michael has acted unfairly in the previous cooperation task*", "*…Peter and Tom intended to punish Michael because of his behavior in the cooperation game*", and "*…it is Michael's own fault how Peter and Tom treat him*", (1 = I do not agree at all, 7 = I agree very much; Cronbach's α = .73). To assess attribution due to a malicious motive, we asked: "*…Peter and Tom are just mean to Michael, absent proper reason.*" "*…Michael cannot be blamed for how Peter and Tom treat him*" "*…Peter and Tom treat Michael the way they do because of the football team he supports.*" (Cronbach's α = .78).

Finally, to check whether participants had understood the presented material correctly, we asked them: "*According to the three persons' choices, which person will probably have to work alone in the subsequent group task?*" and "*Which football teams did the three persons support?*" For research questions unrelated to the present framework, we further assessed dispositional

measures such as observer justice sensitivity (Schmitt, Baumert, Gollwitzer, & Maes, 2010) as well as empathy (Davis, 1983). After providing demographic information, participants were thanked and paid.

Results

Manipulation checks. All participants reported correctly that the target would have to work alone in the subsequent group task. Moreover, 93% of the participants (i.e., 141 out of 152) correctly restated which team each football fan favored.[1]

Dependent variables. A 3 (odd-one-out: target vs. source vs. no one) x 2 (evaluation: target vs. sources) MANOVA on blame, interest in cooperation, anger and sympathy revealed a significant main effect for the evaluation, Wilks' $\lambda = .793$, $F(4, 146) = 9.53$, $p < .001$, $\eta^2 = .21$, reflecting that the target was evaluated differently from the sources. There was no significant main effect of the odd-one-out, $F < 1$. Most importantly, the hypothesized odd-one-out x evaluation interaction was significant, Wilks' $\lambda = .876$, $F(8, 292) = 2.46$, $p = .014$, $\eta^2 = .06$.

To further examine the interaction effect, a series of planned follow-up ANOVAs as well as simple main effects analyses for each of the dependent variables were conducted. The hypothesized interaction effect was significant for every dependent variable (Blame: $F(1, 149) = 8.71$, $p < .001$, $\eta^2 = .11$; Interest in Cooperation: $F(1, 149) = 6.26$, $p = .002$, $\eta^2 = .08$; Anger: $F(1, 149) = 4.48$, $p = .013$, $\eta^2 = .06$; Sympathy: $F(1, 149) = 6.45$, $p = .002$, $\eta^2 = .08$). Consistent with hypotheses, we found that if the target was the odd-one-out, the sources were blamed more, $p = .004$, $d = .65$ and $p = .002$, $d = .71$, and evoked more anger, $p = .033$, $d = .50$ and $p = .068$, $d = .46$, as well as less sympathy, $p = .190$, $d = .38$ and $p = .032$, $d = .49$, compared to the other two conditions where the target was not the odd-one-out. In the condition where the target was the odd-one-out, he was also blamed less, $p = .004$, $d = .63$ and $p = .003$, $d = .70$, and participants

were more sympathetic, $p = .020$, $d = .53$ and $p = .005$, $d = .67$, and more interested in cooperating with him, $p = .009$, $d = .59$ and $p = .002$, $d = .71$.

Within-subject comparisons were conceptually analogous to the between-subjects analyses, insofar that when the target was the odd-one-out, the sources were devalued more than the target, while in the two conditions where the target was not the odd-one-out, the target was devalued more than the sources. For the descriptive statistics and a display of the pattern, see *Table 2* and *Figure 1*.

To investigate the attributions that participants made, we ran a 3 (odd-one-out: target vs. source vs. no one) x 2 (attribution: malicious vs. punitive) ANOVA, with attribution as a repeated measure. There were significant main effects for the odd-one-out, $F(2, 149) = 3.57$, $p = .031$, $\eta^2 = .05$ and for attribution, $F(1, 149) = 11.18$, $p = .001$, $\eta^2 = .07$, which were both qualified by the hypothesized interaction, $F(2, 149) = 22.86$, $p < .001$, $\eta^2 = .24$. Participants attributed ostracism less strongly to a punitive and more strongly to a malicious motive when the target was the odd-one-out compared to the other two conditions, all $p < .001$, smallest $d = .79$. Within-condition comparisons further revealed that when the target was the odd-one-out, participants attributed ostracism more strongly to a punitive motive than to a malicious motive, and vice versa in the other two conditions, all $p < .001$, smallest $d = .47$.

Mediation via attribution. We hypothesized that the effect of the different similarity constellations on the evaluation of the football fans is mediated by different attributions of the situation. More specifically, we assumed that if the target is the odd-one-out, then observers will more strongly attribute ostracism to a malicious motive and less strongly to a punitive motive, and consequently devalue the sources. In contrast, if the target is not the odd-one-out, observers will more strongly attribute ostracism to a punitive motive and less strongly to a malicious

motive, and thus devalue the target. To test the condition of interest (target as odd-one-out)

against the other two conditions, we dummy coded the conditions (odd-one-out: target = 1, other

two conditions = 0) and used this dummy variable as the predictor. As mediators, we

simultaneously tested the influence of the punitive and the malicious attribution measures. As

criteria, we inserted blame, interest in cooperation, anger, and sympathy for the target and the

sources respectively (i.e., eight variables in total). We tested mediation processes with MPLUS

Version 8 (Muthén & Muthén, 1998-2017), using the maximum likelihood estimator with

bootstrapping (5,000 bootstrap estimates). The mediation model was saturated, since we allowed

for all possible relationships between the variables.

The direct path coefficients with 90% confidence intervals are displayed in *Table 3*.

When the target was the odd-one-out, compared to the combined other two conditions,

participants attributed ostracism more strongly to a malicious motive, and as a result of this

attribution, evaluated the target more positively, that is, they blamed the target less, $\beta_{indirect}$ = -

.19, 90% confidence interval (CI) = [-.30; -.09], were more interested in cooperating with him,

$\beta_{indirect}$ = .28, 90% CI = [.18, .38], were less angry, $\beta_{indirect}$ = -.10, 90% CI = [- .17; -.03] and more

sympathetic towards the target, $\beta_{indirect}$ = .25, 90% CI = [.15; .36]. Moreover, participants also

evaluated the sources more negatively, that is, they blamed the sources more, $\beta_{indirect}$ = .31, 90%

CI = [.21; .41], were less interested in cooperating with them, $\beta_{indirect}$ = -.18, 90% CI = [-.28; -

.09], were more angry, $\beta_{indirect}$ = .34, 90% CI = [.22; .46], and less sympathetic towards them,

$\beta_{indirect}$ = -.12, 90% CI = [-.20; -.04], compared to the other two conditions.

Inversely, when the source or no one was the odd-one-out, compared to the target being

the odd-one-out, participants attributed ostracism more strongly to a punitive motive, and as a

result of this attribution, evaluated the target more negatively, that is, they blamed the target

more, $\beta_{indirect}$ = -.14, 90% CI = [-.23; -.08], were less interested in cooperating with him, $\beta_{indirect}$ = .12, 90% CI = [.06; .20], were more angry, $\beta_{indirect}$ = -.24, 90% CI = [-.34; -.15] and less sympathetic towards the target, $\beta_{indirect}$ = .12, 90% CI = [.06; .20]. Moreover, participants evaluated the sources more positively, that is, they were more interested in cooperating with the sources, $\beta_{indirect}$ = -.13, 90% CI = [-.20; -.07], and more sympathetic towards them, $\beta_{indirect}$ = -.20, 90% CI = [-.28; -.13], compared to the other two conditions.

Discussion

Study 2 provides initial support for our hypothesis that in ambiguous situations, observers make use of the social dissimilarity rule to make a moral judgment about the situation: When the target of ostracism was the "odd-one-out" within the group, then participants tended to blame the target less and perceived him as more attractive for cooperation than the sources. Moreover, the target evoked less anger and more sympathy than the sources. Mediation analysis suggests that this might be due to ostracism being attributed to malicious motives of the sources, with malicious motives being presumably fueled by ingroup favoritism. In contrast, the effect dissipated and even partially reversed when the social dissimilarity rule was not applicable because the target was not the "odd-one-out": sources were now seen as more interesting cooperation partners and blamed less; moreover, the sources evoked more sympathy and less anger. Mediation analysis indicated that this might be because ostracism was attributed to a punitive motive: Participants more strongly considered the possibility that the target had acted unfairly during the first task and thus was ostracized by the two sources to punish this norm-violating behavior.

Study 3

Study 3 had two purposes: First, we intended to replicate our findings from Study 2 with

a different sample and in a more controlled setting. Thus, instead of an online sample, we conducted an experiment in the laboratory. Second, we aimed to gather more evidence for the hypothesized process: According to our theoretical assumptions, using the social dissimilarity rule represents a form of heuristic processing or cognitive shortcut, which is used especially in ambiguous situations without sufficient information. Because heuristics can operate with a minimum of cognitive resources, they can be used when individuals do not have the time or motivation to acquire or weigh potentially more complex information (Brewer, 1988; Fiske & Neuberg, 1990; Macrae, et al., 1994; Rudert, Reutner, et al., 2017).

In contrast to the here hypothesized heuristic processing, one could alternatively plead the case that similarity is used as input in systematic processing. In this case, individuals would actively consider similarity and deliberately contemplate its implications. If similarity is used deliberately, however, more processing will be needed compared to heuristic reliance. To put these competing processing assumptions to the test, we imposed cognitive load on participants. A differential pattern of results as a function of cognitive load may be interpreted as an indication of a more deliberative process; a similar pattern of results as an indication of a heuristic process.

Method

Participants and design. Participants were 101 psychology students who participated for course credit. One person did not want his/her data to be analyzed, so that the final sample consisted of 100 participants (87 females, $M_{age} = 22.80$, $SD = 6.41$). Note that sample size was smaller than intended (power analysis based on a three-way interaction effect on a single dependent variable; $f = .25$; power $= .80$; required $n = 136$), due to the fact that not enough participants signed up for the experiment. Participants were randomly assigned to a 2 (odd-one-

out: target vs. source) x 2 (cognitive load: high vs. low) factorial design. We focused on only two odd-one-out conditions because our main interest was in the group for which the social dissimilarity rule would be applicable: the group with the similar sources and the odd-one-out target.

Materials and procedure. The procedure was the same as in Study 2. In addition, before each dependent variable, participants were asked to memorize either an 8-digit number (high cognitive load) or a two-digit number (low cognitive load) in 10 s (Gilbert & Hixon, 1991; Greifeneder & Bless, 2007). Participants recalled the respective numbers directly after answering each dependent variable, and then were provided with the next number to memorize. Assessment in the laboratory ensured that participants could not write down the numbers and thus nullify the cognitive load manipulation. As a manipulation check for cognitive load, we assessed three items such as *"Memorizing the number posed a strong demand on my attention"* (9-point scale, do not agree – agree; Cronbach's α = .86; see Greifeneder & Bless, 2007).

Results

Manipulation checks. Ninety-one percent of the participants reported correctly that the target would have to work alone in the subsequent group task. Moreover, 95% of the participants correctly restated which team each football fan favored. In the low-cognitive load condition only three participants (out of 50) made errors when memorizing the two-digit numbers, while in the high-cognitive load condition, only three participants out of 50 managed to restate all four eight-digit numbers correctly. Importantly, participants in the high load condition reported significantly higher strain than participants in the low load condition, $t(87.14) = -8.33$, $p < .001$, $d = 1.78$ ($M = 5.60$, $SD = 2.18$ vs. $M = 2.47$, $SD = 1.51$).

Dependent variables. A 2 (odd-one-out: target vs. source) x 2 (cognitive load: high vs. low) x 2 (evaluation: target vs. sources) MANOVA on blame, interest in cooperation, anger and sympathy revealed a significant main effect for the evaluation, Wilks' λ = .645, $F(4, 93) = 12.78$, $p < .001$, $\eta^2 = .36$, indicating that in general, the target was evaluated more positively than the sources. Moreover, the hypothesized odd-one-out x evaluation interaction was significant, Wilks' λ = .791, $F(4, 93) = 6.16$, $p = .014$, $\eta^2 = .21$. There was no significant main effect of the odd-one-out, cognitive load or any interaction with cognitive load, smallest $p = .307$.

A series of planned follow-up ANOVAs for each of the dependent variables showed that the hypothesized odd-one-out x evaluation interaction effect was significant for every dependent variable, (Blame: $F(1, 96) = 9.89$, $p = .002$, $\eta^2 = .09$; Interest in Cooperation: $F(1, 96) = 19.92$, $p < .001$, $\eta^2 = .17$; Anger: $F(1, 96) = 6.77$ $p = .011$, $\eta^2 = .07$; Sympathy: $F(1, 96) = 17.26$ $p < .001$, $\eta^2 = .15$. Simple main effects analyses showed that when the target was the odd-one-out, participants blamed the sources more, $p = .001$, $d = .70$, were less interested in cooperating with them, $p = .001$, $d = .66$, and reported more anger, $p = .050$, $d = .39$, and less sympathy for the sources, $p < .001$, $d = .73$, compared to when one of the sources was the odd-one-out. Moreover, participants tended to blame the target less, $p = .015$, $d = .48$, and were more sympathetic, $p = .002$, $d = .62$, and more interested in cooperating with him, $p < .001$, $d = .76$.

Within-subject comparisons were conceptually analogous to the between-subjects analyses, insofar that when the target was the odd-one-out, the sources were devalued more than the target and when one of the sources was the odd-one-out, the target was devalued more than the sources. For the descriptive statistics, see *Table 4*.

A 2 (odd-one-out: target vs. source) x 2 (cognitive load: high vs. low) x 2 (attribution: malicious vs. punitive) ANOVA with attribution as a repeated measure showed a significant

main effect for odd-one-out, $F(1, 96) = 8.62$, $p = .004$, $\eta^2 = .08$ and the attribution, $F(1, 96) =$ 8.97, $p = .003$, $\eta^2 = .09$, which were both qualified by the hypothesized interaction, $F(1, 96) =$ 73.98, $p < .001$, $\eta^2 = .44$. Cognitive load had no significant effect, $F < 1$. Participants attributed ostracism less strongly to a punitive motive, $p < .001$, $d = 1.17$, and more strongly to a malicious motive, $p < .001$, $d = 1.82$, when the target was the odd-one-out compared to when a source was. Within-condition comparisons showed that when the target was the odd-one-out, participants attributed ostracism more strongly to a malicious motive than to a punitive motive, $p < .001$, $d =$ 1.30 and vice versa when a source was the odd-one-out, $p < .001$, $d = .51$.

Mediation via attribution. We ran the same two mediation analyses as in Study 2; see *Table 5* for the direct path coefficients. When the target was the odd-one-out, compared to the source being the odd-one-out, participants attributed ostracism more strongly to a malicious motive, and as a result of this attribution, evaluated the target partly more positively, that is, they were more interested in cooperating with the target, $\beta_{indirect} = .18$, 90% CI = [.04, .31] and were more sympathetic towards the target, $\beta_{indirect} = .20$, 90% CI = [.05; .35]. Moreover, participants evaluated the sources more negatively, that is, they blamed the sources more, $\beta_{indirect} = .18$, 90% CI = [.04; .36], were less interested in cooperating with them, $\beta_{indirect} = -.18$, 90% CI = [-.33; .05], were more angry, $\beta_{indirect} = .44$, 90% CI = [.28; .62], and less sympathetic towards them, $\beta_{indirect} =$ -.26, 90% CI = [-.45; -.09], compared to when a source was the odd-one-out.

Inversely, when a source was the odd-one-out, compared to the target being the odd-one-out, participants attributed ostracism more strongly to a punitive motive, and as a result of this attribution, evaluated the target more negatively, that is, they blamed the target more, $\beta_{indirect} = -$.14, 90% CI = [-.23; -.02], were less interested in cooperating with him, $\beta_{indirect} = .29$, 90% CI = [.20; .40], were more angry, $\beta_{indirect} = -.26$, 90% CI = [-.40; -.16] and less sympathetic towards

the target, $\beta_{indirect}$ = .23, 90% CI = [.13; .34]. Moreover, participants were more sympathetic towards the sources, $\beta_{indirect}$ = -.11, 90% CI = [-.22; -.02], compared to when a source was the odd-one-out.

Discussion

Study 3 replicates the results of Study 2 in a laboratory setting: As in Study 2, the factor odd-one-out affected observers' attributions regarding a social exclusion episode and consequently their evaluation of target and sources. Also, the effect of the odd-one-out on evaluations was mediated via a punitive or a malicious attribution of ostracism. In addition to the factor odd-one-out, we manipulated cognitive load while participants reported their attributions and evaluations. However, we found neither a main effect nor an interaction of cognitive load with odd-one-out on observers' moral judgment. This suggests that individuals are capable of using the social dissimilarity rule under cognitive strain. Albeit being, strictly speaking, a test of the null hypothesis, this finding is very much in line with the hypothesized heuristic nature of the social dissimilarity rule.

It should be noted that statistical power was lower than desired in Study 3. However, the hypothesized odd-one-out x evaluation interaction was still significant. As for a potential effect of cognitive load or any interaction with cognitive load, the descriptive data is not even suggestive of an interaction pattern, rendering it unlikely that sufficient power would have led to significant p-values.

As a potential caveat, it should be noted that cognitive load was manipulated when participants answered the dependent variables, but not when the ostracism situation was presented. Thus, there is the possibility that participants formed their moral judgment immediately when reading and encoding the scenario, that is, before our manipulation of

cognitive load. Although this caveat should not be taken lightly, the alternative methodological approach of manipulating cognitive load before scenario presentation would have been much more problematic, since it might have interfered with participants' reading and thus their general understanding of the scenario, creating different social constructions to begin with.

Study 4

Both Studies 2 and 3 used scenarios. While providing a good initial test of the social dissimilarity rule, the scenario approach has several downsides, such as that it represents a relatively abstract setting in which no information aside from similarity is available. To address this concern, we developed a novel paradigm in which participants observed ostracism in a more realistic setting, namely by reading an excerpt of an alleged chatroom discussion between three discussants. In this excerpt, the contributions and opinions of one person are being ignored. In addition to "cyberostracism" being a common and realistic occurrence (Vorderer & Schneider, 2016; Williams, Cheung, & Choi, 2000; Williams et al., 2002), ostracism in online communication settings, such as chatrooms, mobile phone communication as well as social media platforms, has already been successfully employed in several studies focusing on the target's perspective (Gardner, Pickett, & Brewer, 2000; Smith & Williams, 2004; Wolf et al., 2014). For the present research, a chatroom paradigm further offers the advantage that participants potentially have the chance to form a moral judgment based on multiple cues, such as the contributions of the discussants. These contributions did not clearly point to the reasons for ostracism, however, since we intended to keep the observed situation ambiguous.

Aside from introducing the chatroom paradigm, Study 4 additionally investigates whether the essentiality of the similarity distinction moderates the effect of similarity on moral judgment. This manipulation resulted from the following consideration: not all similarities are equal in the

meaning and importance that people ascribe to them. We wanted to know whether ascriptions of blame and moral judgment require similarity to be deeply rooted, that is, based on something that individuals perceive as an essential distinction, rather than just a superficial coincidence. Building on research about the perceived essentiality of group distinctions, we decided for ethnic background or race as a distinction which is typically perceived as one of the most essential (Haslam, Rothschild, & Ernst, 2000; Prentice & Miller, 2007). As a non-essential distinction, we manipulated the different hairstyles of the discussants and additionally provided participants with the information that the discussants were styled as a part of the experiment and thus did not choose the hairstyles themselves.

Method

Participants and design. Participants were recruited online from Amazon's Mechanical Turk (US Americans only, HIT Approval Rate greater than 97) for a payment of $0.80. Two persons did not want their data to be analyzed. We calculated the sample size such as to detect a medium-sized similarity x essentiality x evaluated person interaction effect on a single dependent variable ($f = .25$, power = .90, required $n = 172$) and recruited 205 participants. Participants were randomly assigned to a 2 (odd-one-out: target vs. source) x 2 (essentiality distinction: essential vs. non-essential) factorial design. Due to a technical error, in one subgroup the agents' portraits were not displayed to participants, so that 14 participants had to be taken out of the sample (1/3 of the essential distinction & target as odd-one-out condition). Thus, the final sample consisted of 191 participants (82 females, 1 other, $M_{age} = 35.91$, $SD = 10.90$). Because of relevance for the present study, we also assessed participants' ethnic background. The majority of participants (79%) self-identified as White/Caucasian, 8% as African American, 7% as Hispanic, 5% as Asian/Pacific Islander and 1% as Native American/Alaskan.

Materials and Pretests.

Development of the chatroom paradigm. Ostracism was introduced by presenting

participants with the logfile of an online chat, see Appendix A. In this chat, three discussants talk

about a group presentation that they have to prepare as part of an experiment. One of the

discussants (the target) is ignored by the other two (the sources) who pay little attention to the

target's suggestions. The names of the discussants are randomized. To allow for variation due to

the similarity manipulation, the content of the chat conversation is rather ambiguous, that is, it

remains unclear whether the target is excluded for her behavior or for other reasons. The chat

was pretested and adapted in a series of pre-studies on both Amazon's Mechanical Turk as well

as in the lab. This was done for two reasons: a) to ensure that participants both clearly understand

that one of the discussants is being excluded and b) to ensure that the chat was sufficiently

ambiguous, so that individuals do not unanimously blame either the target or the sources for the

exclusion. The data and the results for the pretest of the final version of the chat can be accessed

via the Open Science Framework[2].

Facial Portraits. In the present study, the similarity between the actors was manipulated

by means of facial portraits. All discussants were represented by a facial portrait from the MR2

database (Strohminger et al., 2016). For the essential distinction condition, we selected two

female Caucasian, two African American, and two Asian faces. For the non-essential distinction

condition, three Caucasian portraits were manipulated so that the presented persons differed in

their hairstyles (either a bun or a crown braid), see Appendix B for examples of the

manipulation. To make sure that the faces would be perceived as equally similar or non-similar,

we conducted a pretest on Amazon Mechanical Turk with 30 participants (11 female, $M_{age} =$

32.43, $SD = 8.61$). They were shown a set of three portraits and had to decide which one was the

"odd-one-out" as fast as possible. They did this for the four different combinations also used in the main study (two Asians and one of the African American faces, two African American and one of the Caucasian faces, two Caucasian and one of the Asian faces, as well as two persons with a bun and one with a braid hairstyle). The order in which the sets were presented was randomized between participants. More than 90% of all participants selected the intended "odd-one-out", with the exception of the Caucasian/Asian set (68% correct responses). Since this was likely due to one of the Caucasian portraits having lighter hair than the other one, in the main study, the hair colors of the Caucasian portraits were matched.

Procedure. Participants were told that they would read the protocol of an online chat. To introduce the discussants, participants in the *essential distinction condition* saw pictures of three women that differed in ethnic background (Caucasian/Asian, African American/Asian, Caucasian/African American). Importantly, none of the discussants shared an ethnic background with the respective participant; for instance, a Caucasian participant was shown a group that consisted of Asian and African American discussants.

Participants in the *non-essential distinction condition* were shown pictures of three women that differed in their hairstyle. To emphasize the non-essentiality of the discussants' hairstyles, participants were further told that the three discussants got a makeover with a new hairstyle before the picture was taken. Next, participants read the online chat discussion and afterwards evaluated all three discussants on the dependent variables. To assess interest in cooperation, participants were asked: *"How much would you like to work together on a similar project with the following persons?"* (7-point scale, not at all – very much). Anger, sympathy, as well as malicious and punitive intent were assessed as in Studies 1 and 2. To assess whether participants detected ostracism, they were asked: *"Please rate the extent to which each person's*

contributions to the group discussion were acknowledged by the other group members" and

"How much did each person's opinion contribute to the decision making process?" (7-point

scales, not at all – very much). In addition, participants rated each of the three discussants on two

7-point semantic differentials *"excluded – included", "ignored – acknowledged"*. The

evaluations for the sources were averaged across the sources separately for each variable ($\rho = .75$

- .91). Subsequently, participants answered three control questions, e.g., *"Which presentation*

tool did the participants eventually agree on?" and indicated on two 7-point scales how realistic

the chat felt *(very unrealistic – very realistic)* as well as how positive the discussion was

(extremely negative – extremely positive).

Results

     Manipulation checks and control questions. Compared to the sources, participants

reported that the target contributed less to the decision process, $p < .001$, $d = 1.93$ ($M_{Sources} =$

5.94, $SD = 0.96$; $M_{Target} = 2.24$, $SD = 1.35$), that her contributions were acknowledged less, $p <$

.001, $d = 2.23$ ($M_{Sources} = 6.03$, $SD = 1.00$; $M_{Target} = 2.06$, $SD = 1.22$), and that she was more

excluded, $p < .001$, $d = 2.51$ ($M_{Sources} = 6.63$, $SD = 0.74$; $M_{Target} = 2.17$, $SD = 1.47$) and ignored, p

$< .001$, $d = 2.67$ ($M_{Sources} = 6.64$, $SD = 0.72$; $M_{Target} = 2.13$, $SD = 1.41$). The majority of

participants answered the control questions correctly (89%) and felt that the chat was realistic (M

$= 5.29$, $SD = 1.42$) and slightly negative (in comparison to the scale midpoint; $M = 3.55$, $SD =$

1.14).

     Dependent variables. A 2 (odd-one-out: target vs. source) x 2 (essentiality distinction:

essential vs. non-essential) x 2 (discussant: target vs. sources) MANOVA on interest in

cooperation, anger, and sympathy revealed a significant main effect for the discussant, Wilks' λ

$= .609$, $F(3, 185) = 39.62$, $p < .001$, $\eta^2 = .39$, indicating that, in general, participants were more

interested in cooperating with the sources but had more sympathy for the target. Moreover, the similarity x discussant interaction was significant, Wilks' $\lambda = .911$, $F(3, 185) = 6.02$, $p = .001$, $\eta^2 = .09$. There was no significant main effect of similarity, essentiality distinction, or any interaction with essentiality distinction, smallest $p = .102$.

A series of planned follow-up ANOVAs for each of the dependent variables showed that the similarity x discussant interaction was significant for every dependent variable (Interest in Cooperation: $F(1, 187) = 13.76$, $p < .001$, $\eta^2 = .07$, Anger: $F(1, 187) = 14.47$, $p < .001$, $\eta^2 = .07$, Sympathy: $F(1, 187) = 6.37$ $p = .012$, $\eta^2 = .03$. Simple main effects analyses showed that when the target was the odd-one-out, participants were less interested in cooperating with the sources, $p < .001$, $d = .63$, and reported being more angry with the sources, $p = .001$, $d = .51$, than when a source was the odd-one-out. Moreover, participants were more interested in cooperating with the target, $p = .038$, $d = .31$, and felt more sympathetic, $p = .013$, $d = .41$, and less angry towards her, $p = .004$, $d = .41$.

Within-subject comparisons were conceptually analogous to the between-subjects analyses, insofar that when the target was the odd-one-out, the sources were devalued more than the target, while when a source was the odd-one-out, the target was devalued more than the sources. For the descriptive statistics, see *Table 6*.

A 2 (odd-one-out: target vs. source) x 2 (essentiality distinction: essential vs. non-essential) x 2 (attribution: malicious vs. punitive) ANOVA with attribution as a repeated measure showed a significant main effect for odd-one-out, $F(1, 187) = 5.58$, $p = .016$, $\eta^2 = .03$ and attribution, $F(1, 187) = 9.03$, $p = .003$, $\eta^2 = .05$, which were both qualified by the hypothesized interaction, $F(1, 187) = 14.14$, $p < .001$, $\eta^2 = .07$. Essentiality of the group had no significant effects, $Fs < 1$. Participants attributed ostracism less strongly to a punitive motive, $p = .024$, $d =$

.35, and more strongly to a malicious motive, $p < .001$, $d = .66$, when the target was the odd-one-out compared to when a source was. Within-condition comparisons showed that when the target was the odd-one-out, participants attributed ostracism more strongly to a malicious motive than to a punitive motive, $p < .001$, $d = .52$.

      Mediation via attribution. We ran the same two mediation analyses as in Studies 2 and 3; see *Table 7* for the direct path coefficients. When the target was the odd-one-out, compared to the source being the odd-one-out, participants attributed ostracism more strongly to a malicious motive, and as a result of this attribution, evaluated the target partly more positively, that is, they were more interested in cooperating with the target, $\beta_{indirect} = .10$, 90% CI = [.05, .17], were less angry, $\beta_{indirect} = -.07$, 90% CI = [-.11; -.03], and more sympathetic towards the target, $\beta_{indirect} = .18$, 90% CI = [.11; .26]. Moreover, participants evaluated the sources more negatively, that is, they were less interested in cooperating with the sources, $\beta_{indirect} = -.14$, 90% CI = [-.20; .08], were more angry, $\beta_{indirect} = .20$, 90% CI = [.12; .29], and less sympathetic towards them, $\beta_{indirect} = -.13$, 90% CI = [-.19; -.07], compared to when the source was the odd-one-out.

      Inversely, when the source was the odd-one-out, compared to the target being the odd-one-out, participants attributed ostracism more strongly to a punitive motive, and as a result of this attribution, evaluated the target more negatively, that is, they were less interested in cooperating with the target, $\beta_{indirect} = .06$, 90% CI = [.02; .12], were more angry, $\beta_{indirect} = -.10$, 90% CI = [-.16; -.03] and less sympathetic towards the target, $\beta_{indirect} = .05$, 90% CI = [.01; .09]. Moreover, participants were more interested in cooperating with the sources, $\beta_{indirect} = -.06$, 90% CI = [-.11; -.02], less angry, $\beta_{indirect} = .03$, 90% CI = .01; .07], and more sympathetic towards the sources, $\beta_{indirect} = -.07$, 90% CI = [-.11; -.02], compared to when the source was the odd-one-out.

Discussion

Study 4 replicated the results of Studies 2 and 3 with a new paradigm: perceived dissimilarity affected observers' attributions and evaluations of observed ostracism within an online chat. As in the previous studies, the effect of similarity on evaluations was mediated via the attribution of ostracism to a punitive or a malicious motive. In addition to similarity between the group members, we further manipulated the essentiality of the group distinction (ethnic background vs. hairstyle), but found no effect on observers' moral judgment. While one would expect that exclusion of a person with a different skin color might be interpreted as due to racial bias and thus evaluated negatively, exclusion of participants with a different, randomly assigned hairstyle was also attributed to a malicious motive and evaluated negatively. One possible explanation for this is that the social dissimilarity rule is something fundamental which is applied independently of whether differences are essential or not – rather, any hint to dissimilarity is taken as evidence. This would indicate a deeply rooted fairness perspective that no one should be left out just because she/he is dissimilar from the group. Alternatively, one could assume that individuals also have lay theories that a hairstyle might be indicative of culture or other fundamental differences. However, since we provided participants with the information that the discussants were styled as a part of the experiment and thus did not choose the hairstyles themselves, this explanation appears unlikely.

Study 5

In Studies 2-4, we showed that similarity affects observers' moral judgment about a social exclusion episode they had little or very ambiguous information about. In Study 5, we tested whether ambiguity of the situation is in fact a necessary prerequisite for individuals to use the social dissimilarity rule. To do so, we added a beginning to the chat in which the group members introduced themselves and it became obvious whether the target had deviated from

prevailing social norms before being excluded or not. We assumed that if participants received a

more meaningful cue, such as history of the situation, then this would trump the social

dissimilarity rule and thus similarity should not influence participants' moral judgment any

more. In addition, to attest to generalizability of the dissimilarity effect, in Study 5 we

manipulated dissimilarity via the nationality of the group members.

Method

  Participants and design. Participants were recruited online from Prolific Academic (US

Americans only) for a payment of £1.20. We initially calculated the sample size such as to detect

a medium-sized target initial behavior x similarity x evaluated person interaction effect on a

single dependent variable ($f = .25$, power $= .90$, required $n = 206$). The initial sample consisted of

238 participants (116 females, $M_{age} = 32.11$, $SD = 11.16$). After running the study, however, it

became clear that the effect size of the expected three-way interaction was likely smaller than

expected (range of observed fs $= .09$ - .16). Upon suggestion during the review process, we opted

to test whether the three-way interaction shows if statistical power is sufficiently high. We

therefore re-calculated the sample size using an average of the observed effect sizes ($f = .14$,

power $= .80$, required $n = 510$), and collected data of 289 additional participants (128 females,

$M_{age} = 34.68$, $SD = 12.04$) on Prolific Academic. There were no significant differences between

the two samples on any of the dependent variables, all $F < 1$. The final sample thus consisted of

527 participants (244 females, $M_{age} = 33.52$, $SD = 11.71$). Participants were randomly assigned

to a 3 (targets' initial behavior: norm-consistent vs. norm-violating vs. unknown) x 2 (odd-one-

out: target vs. no one) factorial design.

  Materials and Pretest. Similarity was manipulated via the countries that the discussants

came from. To make sure that participants would have no or very few existing associations with

regard to similarities between the countries, we conducted a pretest on Amazon Mechanical Turk with 20 participants (10 female, M_{age} = 30.05, SD = 8.82), testing both a set of the three Baltic countries Latvia, Estonia, and Lithuania, as well as a set of the three Scandinavian countries Sweden, Norway, and Finland. We asked participants to write down their spontaneous associations with each of the countries and then to compare each of the countries with the other two in the respective set. The comparison questions were worded as follows: "*In your opinion, how different from each other are people from X and people from Y?*" and *"How different do you think people from X consider themselves to be from people in Y?"(1 = not different at all; 7 = extremely different).* Descriptive results indicate that participants perceived moderate differences between people from both the three Baltic countries (M = 3.57, SD = 1.18) and the three Scandinavian countries (M = 3.71, SD = 1.07), and assumed that people from these countries would perceive these differences even more strongly (Baltic: M = 4.74, SD = 1.04; Scandinavia: M = 4.34, SD = 1.00). Most importantly, there were no significant differences between the country ratings in the respective sets, all F < 1.08. Because the answers on the open association questions indicated that participants had fewer associations with the Baltic than with the Scandinavian countries, we chose the three Baltic countries for the main study.

In addition to the country of origin, the target's initial behavior was manipulated by presenting participants with different versions of the chat. In the "unknown" condition, participants saw the same chat protocol as in Study 4, which started at a seemingly random point during the discussion. In the two other conditions, the discussants additionally introduced themselves at the beginning of the chat, so that participants would assume they were reading the chatroom discussion from the beginning, see Appendix C. In both conditions, the two sources expressed that they looked forward to the group task. The target then either wrote a friendly

introduction in a similar manner to the sources, thus acting consistent with the social norm of friendliness (norm-consistent condition). In particular, the target wrote: "*I'm [name], pleased to meet and work together with both of you too!*" In the condition supposed to violate the social norm of friendliness (norm-violating condition), the target started with a rude sentence, writing: "*Can we skip the introductions and just get started? I've always hated working in groups.*" After that, the chat continued the same in all three conditions. In a pretest on Prolific Academic with 42 participants (13 female, $M_{age} = 32.50$, $SD = 9.39$), participants saw either the norm-consistent or the norm-violating version of the introduction chat sequence, and were asked about the three discussants' behavior on three 7-point semantic differentials: *improper – proper, rude – friendly, inappropriate – appropriate (*all $\alpha > .97$). Results showed a significant person x condition interaction, $F(2, 38) = 36.87$, $p < .001$, $\eta^2 = .66$. The target's behavior was perceived as more appropriate in the norm-consistent condition than in the norm-violating condition, $F(1,39) = 81.30$, $p < .001$, $\eta^2 = .68$ ($M = 6.67$, $SD = .83$ vs. $M = 2.68$, $SD = 1.82$). In the norm-consistent condition, there were no significant differences between the three discussants, $F < 1$. However, in the norm-violating condition the behavior of the target was seen as less appropriate than the behavior of both sources, $F(2, 38) = 79.72$, $p < .001$, $\eta^2 = .81$.

Procedure. The procedure was similar to Study 4. Participants read the chat that either started without an introduction, or with the norm-consistent, or the norm-violating introduction. During the chat, the respective countries of origin were represented by little flags that were presented next to the names of the three discussants. In the odd-one-out condition, the two sources were from one country and the target from another (e.g., the sources were from Latvia and the target from Estonia). In the condition where no one was the odd-one-out, all discussants

came from the same country. The discussants' names were determined randomly between the three discussants.

Results

Manipulation checks and control questions. Participants reported that the target contributed less to the decision process than the sources, $p < .001$, $d = 1.94$ ($\rho_{Sources} = .76$, $M_{Sources} = 5.97$, $SD = .98$; $M_{Target} = 2.31$, $SD = 1.49$), the target's contributions were acknowledged less, $p < .001$, $d = 2.09$ ($\rho_{Sources} = .80$, $M_{Sources} = 6.02$, $SD = .94$; $M_{Target} = 2.20$, $SD = 1.34$), and that the target was more excluded, $p < .001$, $d = 2.40$ ($\rho_{Sources} = .75$, $M_{Sources} = 6.56$, $SD = 0.73$; $M_{Target} = 2.32$, $SD = 1.56$), and ignored, $p < .001$, $d = 2.34$ ($\rho_{Sources} = .76$, $M_{Sources} = 6.54$, $SD = 0.79$; $M_{Target} = 2.37$, $SD = 1.55$). The majority of participants answered all control questions correctly (92%) and felt that the chat was realistic ($M = 5.07$, $SD = 1.52$), but slightly negative ($M = 3.65$, $SD = 1.31$). On average, participants reported little knowledge about the three Baltic countries, $\alpha = .93$, $M = 1.84$, $SD = 121$.

Dependent variables.

Evaluation. A 3 (targets' initial behavior: norm-consistent vs. norm-violating vs. unknown) x 2 (odd-one-out: target vs. no one) x 2 (discussant: target vs. sources) MANOVA on interest in cooperation, anger, and sympathy revealed a significant main effect for the discussant, Wilks' $\lambda = .587$, $F(3, 519) = 68.45$, $p < .001$, $\eta^2 = .41$. This main effect was qualified by two two-way interactions: odd-one-out x discussant, Wilks' $\lambda = .961$, $F(3, 519) = 7.07$, $p < .001$, $\eta^2 = .04$, and targets' behavior x discussant, Wilks' $\lambda = .802$, $F(6, 1038) = 19.24$, $p < .001$, $\eta^2 = .10$. The three-way interaction was not significant, with $F(6, 1038) = 1.81$, $p = .094$, $\eta^2 = .01$, despite the increase in sample size. Looking at the three variables separately, the three-way interaction for anger was $F(2, 521) = 3.01$, $p = .050$, $\eta^2 = .01$, but was not significant for interest in

cooperation and sympathy, both $F < 1$, $\eta^2 = .00$. Given this finding, which we will discuss further below, we restrict interpretations to the significant two-way interactions.

Specifically, a series of planned follow-up ANOVAs for each of the dependent variables showed that (a) the two-way interaction odd-one-out x discussant was significant for every dependent variable (Interest in Cooperation: $F(1, 521) = 13.60$, $p < .001$, $\eta^2 = .03$, Anger: $F(1, 521) = 4.09$, $p = .027$, $\eta^2 = .01$, Sympathy: $F(1, 521) = 14.13$, $p < .001$, $\eta^2 = .03$) and (b) the two-way interaction target's behavior x discussant effect was significant for every dependent variable (Interest in Cooperation: $F(2, 521) = 52.64$, $p < .001$, $\eta^2 = .17$; Anger: $F(2, 521) = 56.47$, $p < .001$, $\eta^2 = .18$, Sympathy: $F = 37.76$, $p < .001$, $\eta^2 = .13$). We further analyze the two two-way interactions separately in what follows. For the descriptive statistics, see *Table 8 and Figure 2.*

Odd-one-out x discussant. Simple main effects analyses again demonstrated support for the social dissimilarity rule: When the target was the odd-one-out, participants were overall less interested in cooperating with the sources, $p = .009$, $d = .23$, more angry at the sources, $p = .013$, $d = .22$ and reported less sympathy for the sources, $p = .001$, $d = .28$, compared to when there was no odd-one-out. Moreover, participants were overall more interested in cooperating with the target, $p = .001$, $d = .28$, and also reported more sympathy with the target, $p = .002$, $d = .27$, compared to when there was no odd-one-out. Within-subject comparisons were conceptually analogous to the between-subjects analyses, insofar that when the target was the odd-one-out, the sources were devalued more than the target, and when the target was not the odd-one-out, the target was devalued more than the sources (see *Table 8*). However, in both conditions, participants were overall less interested in working together with the target compared to the sources, both $p < .001$, smallest $d = .21$.

Target's behavior x discussant. In the norm-violating condition, participants were more interested in cooperating with the sources, both $p < .001$, smallest $d = .61$ and less angry, both $p < .001$, smallest $d = .60$, and more sympathetic towards the sources, both $p < .001$, smallest $d = .67$, compared to the other two conditions. Moreover, participants were less interested in cooperating with the target, both $p < .001$, $d = .90$, and were more angry, both $p < .001$, smallest $d = .82$, and less sympathetic towards him/her, both $p < .001$, smallest $d = .60$, compared to the other two conditions.

Within-subject comparisons were conceptually analogous to the between-subjects analyses, insofar that when the target behaved norm-consistently or the target's behavior was unknown, participants devalued the sources more than the target (see *Table 8*). In contrast, when the target showed norm-violating behavior at the beginning of the chat, then participants devalued the target more than the sources.

Attribution. A 3 (targets' initial behavior: norm-consistent vs. norm-violating vs. unknown) x 2 (odd-one-out: target vs. no one) x 2 (attribution: malicious vs. punitive) ANOVA with attribution as a repeated measure showed a significant main effect of attribution, $F(1, 521) = 14.05$, $p < .001$, $\eta^2 = .03$ and the target's initial behavior, $F(2, 521) = 4.34$, $p = .014$, $\eta^2 = .02$, an interaction of the target's initial behavior x attribution, $F(2, 521) = 71.71$, $p < .001$, $\eta^2 = .22$, an interaction odd-one-out x attribution, $F(1, 521) = 14.17$, $p < .001$, $\eta^2 = .03$, an interaction target's initial behavior x odd-one-out, $F(2, 521) = 4.04$, $p = .017$, $\eta^2 = .02$, as well as a three-way interaction target's initial behavior x odd-one-out x attribution, $F(2, 521) = 3.53$, $p = .030$, $\eta^2 = .01$. To deconstruct the interactions, we ran follow-up ANOVAs separately for the condition in which the target's behavior was norm-consistent, norm-violating, and unknown. When the target showed *norm-violating behavior* at the beginning of the chat, participants attributed

ostracism more strongly to a punitive motive than to a malicious one, $F(1, 174) = 60.69$, $p <$

.001, $\eta^2 = .26$ ($M_{punitive} = 4.12$, $SD = 1.52$ vs. $M_{malicious} = 2.69$, $SD = 1.27$). In contrast, when the

target showed *norm-consistent behavior* at the beginning of the chat, participants attributed

ostracism more strongly to a malicious motive than to a punitive one, $F(1, 173) = 43.54$, $p <$

.001, $\eta^2 = .20$ ($M_{malicious} = 3.83$, $SD = 1.59$ vs. $M_{punitive} = 2.53$, $SD = 1.32$). The odd-one-out did

not affect individuals' attribution in both conditions, smallest $p = .061$. However, when the

target's previous behavior was unknown, the effect of the attribution, $F(1, 174) = 54.34$, $p <$

.001, $\eta^2 = .24$, was qualified by the significant odd-one-out x attribution interaction, $F(1, 174) =$

17.97, $p < .001$, $\eta^2 = .09$. Participants attributed ostracism more strongly to a malicious motive

when the target was the odd-one-out ($M = 4.50$, $SD = 1.61$) compared to when no one was ($M =$

3.44, $SD = 1.42$), $p < .001$, $d = .70$. In contrast, participants attributed ostracism more strongly to

a punitive motive when there was no odd-one-out ($M = 2.87$, $SD = 1.29$) compared to when the

target was the odd-one-out ($M = 2.38$, $SD = 1.38$), $p = .015$, $d = .37$. Within-condition

comparisons showed that in both conditions, a malicious motive was more strongly assumed than

a punitive motive, $p < .001$, $d = .82$ and $p = .026$, $d = .25$.

Discussion

     In Study 5, we provided participants with information about the history of the exclusion

situation to investigate whether the availability of more diagnostic information would diminish

the use of the social dissimilarity rule. To this end, two thirds of the participants were

additionally provided with a protocol of the chats' opening exchanges, during which the target

acted either consistently with the social norm of being friendly, or violated this norm by acting

rudely. In contrast, one third of participants was not provided with further information about the

chat's beginning sequence (unknown history; same as in Study 4). We observed a strong effect

of knowledge: if the target initially displayed rude behavior (which is inconsistent with the social norm of being friendly), participants attributed ostracism to a punitive motive and sided with the sources. Moreover, in line with our general hypothesis and replicating the previous studies, we observed an effect of the social dissimilarity rule. Contrary to our specific hypothesis in Study 5, however, the three-way interaction did not show, despite sufficient statistical power, suggesting that the effect of the social dissimilarity rule is not moderated by knowledge about the situation.

Ex post we believe that this pattern makes sense. Under the assumption that observers may use simultaneously both knowledge about the situation and the social dissimilarity rule, it is conceivable that information pieces do not interact or qualify each other, but are taken as independent input. In terms of effect sizes, the effect of the target's previous behavior was considerably larger than the effect of the social dissimilarity rule. Thus, it seems that in our chat observer rely on knowledge about the situation's history to a stronger extent than on social dissimilarity information. Potentially, the difference in effect sizes between situational information and social dissimilarity may be a function of the paradigm chosen and is perhaps different when contextual information is even more salient or diagnostic.

Although knowledge about the target's previous behavior did not moderate the effect of the social dissimilarity rule on evaluation, it did so with respect to attributions. In particular, with respect to attributions, social dissimilarity did not affect attributions when participants had information about the target's previous behavior (norm-consistent as well as norm-violating). That the pattern of attributions and evaluations dissociates with respect to the three-way interaction may reflect that attributions are influenced by the available information more directly than evaluations. Again, it is conceivable that this situation changes for paradigms in which contextual information is even more salient or diagnostic.

Methodological Considerations and Empirical Summary

Internal Validity

Manipulation of the independent variables were either pretested and/or tested via manipulation checks in all studies. All continuous manipulation checks were significant. When we used categorical manipulation checks, tests of hypotheses were run with and without the participants who incorrectly answered a manipulation check, which changed neither the significance levels nor the observed patterns. We tested our hypotheses using two different paradigms (scenario and chat paradigm); moreover, in Study 1, we asked people to recall their own observed ostracism experiences to build our studies on peoples' real-life considerations regarding observed ostracism episodes.

Sample Sizes and Diversity

The required sample size was calculated for all studies using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007), opting for a power of at least .80 or higher. Except for Study 3 (the laboratory study), the required n was obtained for all studies. We assessed participants' gender and age in all studies. In the online studies, the gender distribution turned out to be balanced, and the majority of participants were aged between 20 and 40 years. This is in line with samples that can be expected from MTurk and Prolific Academic, which have been found to be more diverse than standard Internet or student samples (Buhrmester, Kwang, & Gosling, 2011; Peer, Brandimarte, Samat, & Acquisti, 2017). While Study 3, the laboratory study, features a much more specific sample (mostly female, Swiss psychology students in their early twenties), we find the same basic effect than in the more diverse, US online samples. Thus, we would expect the effect to generalize across samples, at least within Western cultures.

Mini-Meta Analyses

To gauge the findings' robustness, we conducted two separate meta-analyses on Studies 2-5, testing the effect of the odd-one-out (a) on the evaluation of the target and (b) on the evaluation of the sources. We averaged the effect sizes for the dependent variables assessing evaluation (blame, interest in cooperation, anger, sympathy) to compute one average effect size per study for both the target and the sources. Following the approach suggested by Goh, Hall, and Rosenthal (2016), we used fixed effects in which the mean effect size was weighted by sample size. As for (a), the evaluation of the target, the effect of the odd-one-out was significant, $M\ r = .16, Z = 5.46, p < .001$, two-tailed, such that ostracizing an odd-one-out resulted in observers evaluating the ostracized target more positively. As for (b), the evaluation of the sources, the effect was significant, $M\ r = .18, Z = 5.95, p < .001$, two-tailed, such that ostracizing an odd-one-out resulted in observers evaluating the ostracizing sources more negatively. A fully random effects test of the overall effect was also significant for both the evaluation of the target, $M\ r = .21, t(3)= 4.91, p = .016$, and the evaluation of the sources, $M\ r = .23, t(3)= 5.93, p = .010$.

General Discussion

Observers of ostracism episodes face a critical, yet difficult task to regulate their own behavior: They may wish to decide whether exclusion of the target by the group is justified or not. In the present research, we demonstrate in five studies that ostracism that is attributed to a malicious motive of the sources (such as ingroup favoritism) is perceived as unjustified. In contrast, ostracism that is attributed to a punitive motive appears more justified to observers. We further suggest and empirically substantiate that in ambiguous situations with little information, observers make use of situational cues to decide whether to attribute ostracism to a malicious or a punitive motive. One situational cue that individuals may draw on is social dissimilarity, that

is, the extent to which an ostracized target is similar to the ostracizing sources. This dissimilarity may be a function of values or attitudes; but crucially, it may also be a function of much more easily observable cues that are quickly processed, such as team labels or hairstyles. Four studies provide empirical evidence that when two similar sources exclude a dissimilar target (i.e., the odd-one-out), individuals assume a malicious motive of the sources and report more anger and less interest in cooperation with the sources as well as more sympathy for the target, compared to other group constellations. In contrast, if the target is similar to the sources, individuals more likely assume a punitive motive, which results in less sympathy and interest to cooperate with the target as well as more anger. This general pattern of findings is robust across two different paradigms: a scenario in Studies 2 and 3, and the observation of a group chat in Studies 4 and 5. We further demonstrate that the effect of the social dissimilarity rule is neither reduced by a cognitive capacity manipulation (Study 3), nor by perceived essentialism of the group distinction (Study 4), which stresses our notion that the social dissimilarity rule is general, frugal, and processed with ease. While dissimilarity resulting from easily observable cues allows for quick processing, these cues may not constitute very valid information. As a result, we predicted that the social dissimilarity rule (based on easily observable cues) is relied on primarily in ambiguous situations, that is, when more diagnostic information is lacking. Consistent with this reasoning, participants in Study 5 primarily based their judgement on more diagnostic information if such information became available, such as whether the target had previously violated a social norm by acting rudely. Nevertheless, social dissimilarity still continued to affect observer's moral judgment even in situations that provide a richer background.

Observers' Attributions of Ostracism

The present research adds to a body of literature which emphasizes the importance of individuals' interpretations and attributions of ostracism situations (Arpin, et al., 2017; Rudert & Greifeneder, 2016; Rudert, Reutner, et al., 2017; Tuscherer et al., 2015; Wesselmann, et al., 2014; Wesselmann, Wirth, et al., 2013). Earlier studies more or less explicitly implied that observers of ostracism unconditionally tend to empathize with the targets and "feel their pain" (Legate, DeHaan, Weinstein, & Ryan, 2013; Masten, et al., 2011; Wesselmann, et al., 2009; Will, et al., 2013). However, the present research indicates that observers instead consider whether ostracism represents a norm violation or a rightful punishment and make a moral judgment based on the available cues of the situation. One reason why earlier research primarily observed empathizing with the target may be the methodology employed: In the majority of previous observer studies, participants observed or participated in a stand-alone Cyberball game with players that lack a common history. In this situation, individuals have no reason to assume that the target has committed a norm violation that the other players are aware of. In contrast, in the paradigms used in the present research, observers knew that there was some previous common history, and that there was the possibility that the target had deviated from the norm before by being uncooperative and/or rude.

By stipulating that observers do not unconditionally side with the target, the present research also highlights the important role of *motives* (here: punitive and malicious), that has been discussed in the beginning of ostracism research (Sommer, Williams, Ciarocco, & Baumeister, 2001; Williams, 2002). However, research on motives of the sources has not received much attention in recent years, possibly because motives seemed to be of minor importance for the *targets'* reactions and well-being and ostracism even hurts when it is unintentional (Williams, 2009; Zadro, Williams, & Richardson, 2004). Yet, as the present

research demonstrates, from the *observers'* perspectives, it is crucial to take into account which attributions observers make regarding the sources' motives, especially whether observers assume an underlying punitive or a more selfish, malicious motive of the ostracizing sources.

Development of moral attributions for ostracism

The present findings provide evidence that observers dislike sources who ostracize the "odd-one-out," because they attribute ostracism to a malicious motive such as ingroup favoritism and discrimination. However, the picture changes when the target has violated social norms: such punitive ostracism is perceived as justified and acceptable. One question one may reasonably ask is from where these consistently observed attributions result. On a speculative note, we assume that they are learned during an individual's ontogenesis, and are deeply ingrained in cultural knowledge that is socially transmitted from the earliest days on. For instance, an important lesson taught to children in kindergarten and primary school is that it is wrong to ostracize others just because they are different. Consider famous children's books such as Hans Christian Andersen's story of the Ugly Duckling or Johnny Mark's popular song about Rudolph, the red-nosed reindeer, who is not allowed to "join in any reindeer games" because of his brightly shining nose. Reversely, at least in the culture the authors were brought up, children in kindergarten and school are often told to "just ignore" a bully instead of physically aggressing. In doing so, they learn that ostracism is not only socially acceptable, but an appropriate way to deal with troublesome or deviant others. Consistent with this perspective, punitive or defensive ostracism are often perceived to be socially acceptable alternatives to direct or physical aggression (Robinson, O'Reilly, & Wang, 2013).

Prevalence of the Social Dissimilarity Rule

We observed a robust effect of the social dissimilarity rule in situations that lack diagnostic information. How prevalent are situations such as these? We think they are plenty. Consider all kinds of transition phases, such as when an individual enters university or starts a new job, or when individuals move to another town or neighborhood. Moreover, consider all kinds of public events, such as when individuals go to the movies or for a walk through the city. In these situations, observers likely lack diagnostic information of prior social context. However, making a moral judgment about ostracism might still be of utmost relevance: While an incomer may not wish to familiarize with a group of co-workers who ostracize based on race, she might also be glad to stay away from the self-absorbed, rude colleague who is ostracized by everyone else. To the extent that many situations are characterized by incomplete prior knowledge, we thus speculate that the social dissimilarity rule is relied on quite frequently.

Alternative Explanation: Differences in Perceived Hurt

We have argued that the evaluation of targets and observers is mainly driven by observers' moral judgment of the ostracism episode and the attributions on the sources' underlying motives. This relation was also emphasized by the analyses indicating that the effect of dissimilarity on evaluation was mediated via attributions to different motives. In addition, one could speculate about whether the observed effect could also be driven by the observers' assessment of how aversive and hurtful ostracism may be for the ostracized target. In other words, observers might experience more empathy if they assume that ostracism felt more hurtful for the target. Such differences in felt hurt could presumably arise as a function of group membership. However, given that observers would presumably assume that ostracism by the ingroup hurts the target more than ostracism by the outgroup, according to this explanation we should have found the exact opposite effect of the observed evidence, namely a more positive

evaluation of the target when she/he is ostracized by her/his ingroup (similarity) compared to her/his outgroup (dissimilarity). The differences-in-hurt alternative explanation is also rather difficult to reconcile with the observed evidence on mediation via motive attribution.

As a potential caveat, it should be mentioned that all mediation analyses were performed on cross-sectional data. Conclusions about causality can be therefore be tentative only. Nevertheless, it should be noted that we specified the mediation models against the background of the a priori theoretical assumption that attributions precede evaluations. Although it is possible to imagine contexts in which the reversed causal relation is true as well (i.e., evaluations affecting attributions), such a relation appears less plausible from a theoretical standpoint, especially if participants have no previous information or opinion about the persons they are about to evaluate, as it was the case in our experiments. Note that the present assumption about a causal link from attributions to evaluations is consistent with other research (e.g., Arpin, et al., 2017; Chatman & Von Hippel, 2001).

Consequences of Moral Judgments

So far, we have tacitly assumed that observers' moral judgments matter. We wish to close with a short speculation on the downstream consequences of observers' moral judgments to bolster this assumption. First, to the extent that observers' judgments are known to the group members, observers' judgments will likely matter for the targets' feelings and behavior (as the observer may be perceived as an ally or not), as well as for the sources' feelings and behavior (as they may perceive approval or disapproval). As such, the mere judgments may prove quite consequential. For instance, a child that is ostracized by some of his peers may take comfort in knowing if either other peers, or her teachers, signal that they do not approve of how the child is treated. Second, as a further downstream consequence, observers' moral judgments may impact

the observers' behavior, such as bystander intervention. Bystander intervention is likely

influenced by many variables and considerations other than moral judgment, such as whether the

observer deems that it is appropriate or necessary to intervene (Chekroun & Brauer, 2002;

Fischer et al., 2011; Latané & Darley, 1969). Nevertheless, moral judgments are one predictor of

helping behavior and may therefore play a critical role.

We have demonstrated that when making moral judgements about ostracism episodes,

observers may rely on the social dissimilarity rule. These judgments may prove right or wrong

depending on the cue validity between social dissimilarity and exclusion cause. We have no data

with respect to cue validity. Nevertheless, it appears commendable to briefly discuss how the cue

may be wrong, and ways to reduce the rule's impact in applied settings. For instance, if sources

exclude a target due to a dissimilarity that is not easily observable, observers might fail to protect

an innocent target because they do not detect that the target is the odd-one-out, and thus assume

that the target must have done something wrong. Alternatively, the target might be dissimilar

from the sources by coincidence, although the dissimilarity is not the underlying reason for why

the target is ostracized. In this case, observers may misattribute ostracism to a malicious motive

and side with a target who has actually hurt the group and instead blame the group that had only

tried to protect itself from a harmful member.

A potential solution might be to provide or encourage individuals to seek more

information before making a moral judgment, consistent with the findings of Study 5. Against

this background, individuals who have to deal with matters of ostracism and exclusion in

professional contexts, such as teachers, counselors, conflict mediators, and HR employees,

should ideally be trained in carefully questioning the entire story behind an ostracism episode, in

order to reduce the possibility that their judgment becomes distorted by stereotypes about ostracism situations and ostracizers.

Conclusion

Taken together, the presented five studies show that observers' moral judgment of ambiguous ostracism situations is influenced by the dissimilarity between the observed group members. Especially when a dissimilar person is ostracized by a homogeneous group, observers tend to attribute ostracism to a malicious motive and blame as well as devalue the ostracizing sources for the situation. In contrast, if the target is not apparently dissimilar from the rest of the group, observers are more likely to attribute ostracism to a punitive motive and blame and devalue the target instead. The social dissimilarity rule proved robust across several paradigms and situational variations. Together, the studies provide valuable insight into how observers make moral judgments about ostracism.

References

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

Brewer, M. B. (1988). A dual process model of impression formation. In R. S. Wyer & T. K.

Srull (Eds.), *Advances in Social Cognition.* Hillsdale, NJ: Erlbaum.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source

of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*(1), 3-5.

doi: 10.1177/1745691610393980

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

Crandall, C. S., Eshleman, A., & O'Brien, L. (2002). Social norms and the expression and

suppression of prejudice: The struggle for internalization. *Journal of Personality and

Social Psychology, 82*(3), 359-378. doi: 10.1037/0022-3514.82.3.359

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a

multidimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113-

126. doi: 10.1037/0022-3514.44.1.113

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

Eck, J., Schoel, C., & Greifeneder, R. (2016). Coping with or buffering against the negative

  impact of social exclusion on basic needs: A review of strategies. In P. Riva & J. Eck

  (Eds.), *Social exclusion: Psychological approaches to understanding and reducing its*

  impact (pp. 227-249). New York, NY: Springer.

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

Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., . . .

  Kainbacher, M. (2011). The bystander-effect: a meta-analytic review on bystander

  intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin,*

  137(4), 517-537. doi: 10.1037/a0023304

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-

  based to individuating processes: Influences of information and motivation on attention

  and interpretation. *Advances in Experimental Social Psychology, 23*, 1-74. doi:

  10.1016/S0065-2601(08)60317-2

Gardner, W. L., Pickett, C. L., & Brewer, M. B. (2000). Social exclusion and selective memory:

  How the need to belong influences memory for social events. *Personality and Social*

  Psychology Bulletin, 26(4), 486-496. doi: 10.1177/0146167200266007

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of

    stereotypic beliefs. *Journal of Personality and Social Psychology, 60*(4), 509-517. doi:

    10.1037/0022-3514.60.4.509

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some

    arguments on why and a primer on how. *Social and Personality Psychology Compass,*

    10(10), 535-549. doi: 10.1111/spc3.12267

Gooley, S. L., Zadro, L., Williams, L. A., Svetieva, E., & Gonsalkorale, K. (2015). Ostracizing

    for a Reason: A novel source paradigm for examining the nature and consequences of

    motivated ostracism. *The Journal of Social Psychology, 155*(5), 410-431. doi:

    10.1080/00224545.2015.1060933

Greifeneder, R., & Bless, H. (2007). Relying on accessible content versus accessibility

    experiences: The case of processing capacity. *Social Cognition, 25*(6), 853-881. doi:

    10.1521/soco.2007.25.6.853

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories.

    British Journal of Social Psychology, 39(1), 113-127. doi: 10.1348/014466600164363

Inman, M. L., & Baron, R. S. (1996). Influence of prototypes on perceptions of prejudice.

    Journal of Personality and Social Psychology, 70(4), 727-739. doi: 10.1037/0022-

    3514.70.4.727

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

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics, 33*(1), 159-174. doi: 10.2307/2529310

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

Legate, N., DeHaan, C. R., Weinstein, N., & Ryan, R. M. (2013). Hurting you hurts me too: The psychological costs of complying with ostracism. *Psychological Science, 24*(4), 583-588. doi: 10.1177/0956797612457951

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology, 66*(1), 37-47. doi: 10.1037/0022-3514.66.1.37

Marques, J. M., & Paez, D. (1994). The 'black sheep effect': Social categorization, rejection of ingroup deviates, and perception of group variability. *European Review of Social Psychology, 5*(1), 37-68. doi: 10.1080/14792779543000011

Marques, J. M., Yzerbyt, V. Y., & Leyens, J. P. (1988). The "black sheep effect": Extremity of judgments towards ingroup members as a function of group identification. *European Journal of Social Psychology, 18*(1), 1-16. doi: 10.1002/ejsp.2420180102

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

McKown, C., & Weinstein, R. S. (2003). The development and consequences of stereotype consciousness in middle childhood. *Child Development, 74*(2), 498-515. doi: 10.1111/1467-8624.7402012

Moss-Racusin, C., Phelan, J., & Rudman, L. (2010). "I'm not prejudiced, but...": Compensatory egalitarianism in the 2008 Democratic Presidential Primary. *Political Psychology, 31*(4), 543-561. doi: 10.1111/j.1467-9221.2010.00773.x

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide. Eighth Edition*. Los Angeles, CA: Muthén & Muthén.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70*, 153-163. doi: 10.1016/j.jesp.2017.01.006

Prentice, D. A., & Miller, D. T. (2007). Psychological essentialism of human categories. *Current Directions in Psychological Science, 16*(4), 202-206. doi: 10.1111/j.1467-8721.2007.00504.x

Robinson, S. L., O'Reilly, J., & Wang, W. (2013). Invisible at work: An integrated model of workplace ostracism. *Journal of Management, 39*(1), 203-231. doi: 10.1177/0149206312466141

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., 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., Janke, S., & Greifeneder, R. (2017). Under threat by popular vote: German-speaking immigrants' affect and cognitions following the Swiss vote against mass immigration. *PLOS ONE, 12*(4). doi: 10.1371/journal.pone.0175896.

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

Schmitt, M., Baumert, A., Gollwitzer, M., & Maes, J. (2010). The Justice Sensitivity Inventory: Factorial validity, location in the personality facet space, demographic pattern, and normative data. *Social Justice Research, 23*(2-3), 211-238. doi: 10.1007/s11211-010-0115-2

Smith, A., & Williams, K. D. (2004). R u there? Ostracism by cell phone text messages. *Group Dynamics: Theory, Research, and Practice, 8*(4), 291-301. doi: 10.1037/1089-2699.8.4.291

Sommer, K. L., Williams, K. D., Ciarocco, N. J., & Baumeister, R. F. (2001). When silence speaks louder than words: Explorations into the intrapsychic and interpersonal consequences of social ostracism. *Basic and Applied Social Psychology, 23*(4), 225-243. doi: 10.1207/s15324834basp2304_1

Sommers, S. R., & Norton, M. I. (2006). Lay theories about White racists: What constitutes racism (and what doesn't). *Group Processes & Intergroup Relations, 9*(1), 117-138. doi: 10.1177/1368430206059881

Strohminger, N., Gray, K., Chituc, V., Heffner, J., Schein, C., & Heagins, T. B. (2016). The MR2: A multi-racial mega-resolution database of facial stimuli. *Behavior Research Methods, 48*(3), 1197-1204. doi: 10.3758/s13428-015-0641-9

Teng, F., & Chen, Z. (2012). Does social support reduce distress caused by ostracism? It depends on the level of one's self-esteem. *Journal of Experimental Social Psychology, 48*(5), 1192-1195. doi: 10.1016/j.jesp.2012.03.014

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

Vivian, J. E., & Berkowitz, N. H. (1992). Anticipated bias from an outgroup: An attributional analysis. *European Journal of Social Psychology, 22*(4), 415-424. doi: 10.1002/ejsp.2420220410

Vorderer, P., & Schneider, F. (2016). Social media and ostracism. In K. D. Williams & S. A. Nida (Eds.), *Ostracism, exclusion, and rejection* (pp. 240-257). New York, NY: Psychology Press.

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

Wesselmann, E. D., Williams, K. D., & Wirth, J. H. (2014). Ostracizing group members who can (or cannot) control being burdensome. *Human Ethology Bulletin, 29*(2), 82-103.

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

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. (2002). *Ostracism: The power of silence*. New York, NY: Guilford Press.

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., Govan, C. L., Croker, V., Tynan, D., Cruickshank, M., & Lam, A. (2002). Investigations into differences between social- and cyberostracism. *Group Dynamics: Theory, Research, and Practice, 6*(1), 65-77. doi: 10.1037/1089-2699.6.1.65

Wolf, W., Levordashka, A., Ruff, J. R., Kraaijeveld, S., Lueckmann, J.-M., & Williams, K. D. (2014). Ostracism Online: A social media ostracism paradigm. *Behavior Research Methods*, 1-13. doi: 10.3758/s13428-014-0475-x

Zadro, L., Williams, K. D., & Richardson, R. (2004). How low can you go? Ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem, and meaningful existence. *Journal of Experimental Social Psychology, 40*(4), 560-567. doi: 10.1016/j.jesp.2003.11.006

Zhou, X., & Gao, D.-G. (2008). Social support and money as pain management mechanisms. *Psychological Inquiry, 19*(3-4), 127-144. doi: 10.1080/10478400802587679

Footnotes

[1] In this study as well as all subsequent studies, excluding participants from the analyses who failed to answer one or more control questions changed neither the pattern of results nor the significance levels within the MANOVA. Thus, all analyses throughout the manuscript are based on the full number of participants.

[2] https://osf.io/6g9vk/?view_only=4f52a2d1f5e149b5a220633a3042f9a4

Table 1

Results for Study 1 b

Dependent Variable	Situation 1 (Norm-violating Target)	Situation 2 (Norm-violating Target)	Situation 3: (Dissimilar Target)	Situation 4: (Dissimilar Target)
Ostracism Justified	4.80 (1.50)	4.28 (1.49)	1.30 (1.04)	1.58 (.99)
Anger (at Target)	3.18 (1.95)	2.92 (1.74)	2.22 (2.19)	1.98 (1.78)
Sympathy (for Target)	3.64 (1.94)	4.04 (1.64)	6.36 (.92)	6.22 (1.06)
Anger (Situation)	3.00 (1.80)	3.04 (1.76)	5.62 (1.63)	4.82 (1.76)
Sadness (Situation)	3.50 (1.95)	3.86 (1.83)	6.04 (1.23)	5.84 (1.27)

Note. Means (and standard deviations) as a function of the four experimental conditions (situations).

Table 2

Results for Study 2

Dependent Variable	Repeated Measure	Target as odd-one-out	Source as odd-one-out	No odd-one-out
Blame	Target	3.48 [a] (2.17)	4.94 [b] (2.47)	5.00 [b] (2.20)
	Sources	4.80 [c] (1.76)	3.58 [d] (2.00)	3.49 [d] (1.92)
Interest in Cooperation	Target	4.50 [a] (2.00)	3.29 [b] (2.13)	3.08 [b] (1.99)
	Sources	3.57 [c] (1.83)	4.34 [c] (1.63)	4.28 [c] (1.69)
Anger	Target	2.15 [a] (1.76)	2.67 [a] (2.00)	2.92 [a] (1.94)
	Sources	3.32 [b] (1.83)	2.43 [a] (1.72)	2.51 [ab] (1.70)
Sympathy	Target	4.60 [a] (1.89)	3.56 [b] (2.03)	3.35 [b] (1.84)
	Sources	2.81 [c] (1.90)	3.49 [bc] (1.66)	3.77 [b] (2.01)
Motive Strength	Punitive	3.17 [a] (1.51)	4.35 [b] (1.48)	4.48 [b] (1.56)
	Malicious	4.48 [c] (1.60)	2.87 [d] (1.50)	2.48 [d] (1.20)

Note. Means (and standard deviations) as a function of the three 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

Direct Path Coefficients of the Mediation Model in Study 2

	β	*90% CI*
Condition → Punitive Attribution	-.37	-.49; -.24
Condition → Malicious Attribution	.51	.40; .61
Condition → Blame Target	.03	-.09; .16
Condition → Cooperation with Target	-.11	-.21; -.00
Condition → Anger at Target	.18	.08; .27
Condition → Sympathy with Target	-.10	-.21; .02
Condition → Blame Sources	-.03	-.16; .09
Condition → Cooperation with Sources	.11	-.02; .24
Condition → Anger at Sources	-.14	-.26; -.01
Condition → Sympathy with Sources	.11	.00; .22
Punitive Attribution → Blame Target	.39	.25; .54
Punitive Attribution → Cooperation with Target	-.33	-.47; -.21
Punitive Attribution → Anger at Target	.65	.53; .77
Punitive Attribution → Sympathy with Target	-.32	-.49; -.18
Punitive Attribution → Blame Sources	-.09	-.24; .06
Punitive Attribution → Cooperation with Sources	.37	.23; .51
Punitive Attribution → Anger at Sources	-.08	-.24; .07
Punitive Attribution → Sympathy with Sources	.55	.40; .68
Malicious Attribution → Blame Target	-.37	-.54; -.19
Malicious Attribution → Cooperation with Target	.55	.40; .68
Malicious Attribution → Anger at Target	-.19	-.31; -.06
Malicious Attribution → Sympathy with Target	.50	.32; .65
Malicious Attribution → Blame Sources	.60	.45; .74
Malicious Attribution → Cooperation with Sources	-.35	-.50; -.19
Malicious Attribution → Anger at Sources	.66	.48; .82
Malicious Attribution → Sympathy with Sources	-.23	-.37; -.09

Note. The non-directional path between malicious and punitive attribution was $r = -.61$, 90% CI = [-.71; -51]. Correlations between the dependent variables are not displayed.

Table 4

Results for Study 3

Dependent Variable	Repeated Measure	Low Load		High Load	
		Target as odd-one-out	*Source as odd-one-out*	*Target as odd-one-out*	*Source as odd-one-out*
Blame	Target	4.19 [a] (2.26)	4.75 [ab] (2.17)	2.92 [c] (2.00)	4.60 [a] (2.55)
	Sources	4.67 [ad] (1.67)	3.63 [bc] (1.76)	5.22 [d] (1.44)	3.92 [ac] (1.70)
Interest in Cooperation	Target	5.27 [a] (1.59)	4.21 [bc] (2.15)	6.08 [a] (.95)	4.40 [c] (2.16)
	Sources	3.42 [c] (1.63)	4.65 [b] (1.46)	3.42 [c] (1.61)	4.16 [bc] (1.25)
Anger	Target	1.77 [a] (1.37)	2.37 [a] (1.50)	1.68 [a] (1.28)	2.12 [a] (1.64)
	Sources	3.60 [b] (1.75)	2.94 [ab] (1.61)	4.08 [b] (1.99)	3.34 [b] (1.66)
Sympathy	Target	5.31 [a] (1.35)	4.21 [c] (1.82)	5.52 [a] (1.26)	4.60 [c] (1.94)
	Sources	2.94 [b] (1.37)	4.02 [cd] (1.20)	2.88 [b] (1.16)	3.64 [d] (1.29)
Motive Strength	Punitive	2.76 [a] (1.36)	4.21 [b] (1.27)	2.41 [a] (.87)	4.04 [b] (1.61)
	Malicious	5.18 [b] (1.33)	2.74 [a] (1.23)	5.16 [b] (1.10)	3.01 [a] (1.38)

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 bonferroni-corrected 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 5

Direct Path Coefficients of the Mediation Model in Study 3

	β	90% CI
Condition → Punitive Attribution	-.51	-.64; -.37
Condition → Malicious Attribution	.68	.59; .76
Condition → Blame Target	-.04	-.26; .20
Condition → Cooperation with Target	-.11	-.28; .07
Condition → Anger at Target	.13	-.05; .30
Condition → Sympathy with Target	-.13	-.32; .07
Condition → Blame Sources	.04	-.16; .25
Condition → Cooperation with Sources	-.05	-.23; .14
Condition → Anger at Sources	-.20	-.41; .01
Condition → Sympathy with Sources	.02	-.17; .24
Punitive Attribution → Blame Target	.27	.04; .46
Punitive Attribution → Cooperation with Target	-.57	-.70; -.42
Punitive Attribution → Anger at Target	.52	.35; .68
Punitive Attribution → Sympathy with Target	-.44	-.59; -.28
Punitive Attribution → Blame Sources	-.21	-.41; .03
Punitive Attribution → Cooperation with Sources	.16	.05; .35
Punitive Attribution → Anger at Sources	.09	-.08; .26
Punitive Attribution → Sympathy with Sources	.22	.03; .40
Malicious Attribution → Blame Target	-.09	-.35; .13
Malicious Attribution → Cooperation with Target	.26	.05; .45
Malicious Attribution → Anger at Target	-.06	-.30; .18
Malicious Attribution → Sympathy with Target	.30	.07; 50
Malicious Attribution → Blame Sources	.27	.07; .50
Malicious Attribution → Cooperation with Sources	-.27	-.47; -.08
Malicious Attribution → Anger at Sources	.65	.43; .86
Malicious Attribution → Sympathy with Sources	-.38	-.64; -.13

Note. The non-directional path between malicious and punitive attribution was *r* = -.50, 90% CI = [-.63; -36]. Correlations between the dependent variables are not displayed.

Table 6

Results for Study 4.

Dependent Variable	Repeated Measure	Non-essential Group		Essential Group	
		Target as odd-one-out	*Source as odd-one-out*	*Target as odd-one-out*	*Source as odd-one-out*
Interest in Cooperation	Target	3.47 [a] (1.75)	3.00 [a] (1.82)	3.53 [a] (2.00)	2.87 [a] (1.75)
	Sources	3.50 [a] (1.98)	4.58 [b] (1.74)	3.48 [a] (1.80)	4.69 [b] (1.82)
Anger	Target	3.15 [a] (1.93)	3.86 [ac] (1.97)	2.87 [a] (1.59)	3.76 [c] (1.86)
	Sources	4.04 [b] (1.71)	3.23 [a] (1.74)	4.15 [b] (2.13)	3.05 [ac] (2.03)
Sympathy	Target	4.76 [a] (1.83)	4.00 [b] (2.09)	4.40 [a] (1.98)	3.67 [ab] (2.04)
	Sources	3.25 [b] (1.59)	3.78 [b] (1.91)	3.28 [ab] (1.73)	3.76 [ab] (1.90)
Motive Strength	Punitive	2.63 [a] (1.45)	3.15 [a] (1.41)	2.71 [a] (1.47)	3.20 [a] (1.59)
	Malicious	4.05 [b] (1.53)	3.24 [a] (1.53)	4.00 [b] (1.67)	2.81 [a] (1.46)

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 bonferroni-corrected 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 7

Direct Path Coefficients of the Mediation Model in Study 4

	β	90% CI
Condition → Punitive Attribution	-.17	-.29; -.06
Condition → Malicious Attribution	.32	.21; .42
Condition → Cooperation with Target	-.01	-.11; .08
Condition → Anger at Target	-.04	-.12; .05
Condition → Sympathy with Target	-.03	-.11; .04
Condition → Cooperation with Sources	-.10	-.18; -.01
Condition → Anger at Sources	.01	-.07; .09
Condition → Sympathy with Sources	.05	-.04; .14
Punitive Attribution → Cooperation with Target	-.37	-.50; -.23
Punitive Attribution → Anger at Target	.56	.46; .65
Punitive Attribution → Sympathy with Target	-.28	-.38; -.17
Punitive Attribution → Cooperation with Sources	.35	.24; .46
Punitive Attribution → Anger at Sources	-.19	-.30; -.08
Punitive Attribution → Sympathy with Sources	.38	.26; .49
Malicious Attribution → Cooperation with Target	.33	.18; .48
Malicious Attribution → Anger at Target	-.21	-.33; -.10
Malicious Attribution → Sympathy with Target	.57	.48; .68
Malicious Attribution → Cooperation with Sources	-.44	-.55; -.33
Malicious Attribution → Anger at Sources	.63	.53; .73
Malicious Attribution → Sympathy with Sources	-.40	-.51; -.29

Note. The non-directional path between malicious and punitive attribution was $r = -.58$, 90% CI = [-.66; -50]. Correlations between the dependent variables are not displayed.

Table 8

Results for Study 5.

Dependent Variable	Repeated Measure	Norm-violating Target		Norm-consistent Target		Unknown Target Behavior	
		Target as odd-one-out	No odd-one-out	Target as odd-one-out	No odd-one-out	Target as odd-one-out	No odd-one-out
Interest in Cooperation	Target	2.35 [a] (1.61)	1.93 [a] (1.25)	3.83 [b] (1.80)	3.30 [c] (1.60)	3.81 [b] (1.64)	3.33 [bc] (1.77)
	Sources	4.94 [e] (1.73)	5.03 [e] (1.68)	3.65 [b] (1.94)	4.20 [d] (1.61)	3.42 [b] (1.81)	4.01 [d] (1.78)
Anger	Target	4.88 [a] (1.46)	4.60 [a] (1.79)	3.31 [cb] (1.66)	3.45 [c] (1.71)	2.95 [b] (1.64)	3.44 [c] (1.63)
	Sources	2.69 [b] (1.70)	2.65 [b] (1.78)	3.99 [a] (1.83)	3.49 [ac] (1.88)	4.15 [a] (1.79)	3.52 [c] (1.74)
Sympathy	Target	3.76 [a] (1.67)	3.33 [a] (1.86)	4.84 [bd] (1.85)	4.41 [b] (1.84)	5.01 [d] (1.59)	4.48 [b] (1.72)
	Sources	4.53 [b] (1.61)	4.70 [b] (1.74)	3.22 [c] (1.59)	3.77 [ae] (1.61)	3.20 [c] (1.66)	3.83 [e] (1.50)
Motive Strength	Punitive	4.01 [a] (1.44)	4.22 [a] (1.58)	2.43 [c] (1.33)	2.63 [cd] (1.30)	2.38 [c] (1.38)	2.87 [d] (1.29)
	Malicious	2.66 [c] (1.20)	2.71 [c] (1.34)	4.09 [a] (1.77)	3.55 [b] (1.33)	4.50 [a] (1.61)	3.44 [b] (1.42)

Note. Means (and standard deviations) as a function of the six experimental conditions, separately for targets and sources of ostracism. The values in brackets are the standard deviations. The letters a - d represent bonferroni-corrected 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.

Appendix A
Chat Room Paradigm (Studies 4 and 5)

In this study we investigate how people form first impressions of other people in the absence of social interaction.

In the following activity, we will show you the history from a recent online chat, which we took from another experiment with university students. The chat history contains text from a conversation among three students who are preparing a short group presentation on the topic of "team-building". The three participants of the chat have been nicknamed *Cube*, *Soda*, and *Jazz*.

Cube, *Soda*, and *Jazz* are students with different majors, but from the same university. These students did not know each other before the experiment, and were randomly assigned into the chat group as a part of the study they were participating in.

During their conversation, *Cube*, *Soda*, and *Jazz* first introduced themselves and then discussed about the contents and the sequences of their presentation.

As in face-to-face conversations, online group discussion can be more or less harmonious.

On the next page we will show you an excerpt of the online group discussion which occurred between Cube, Soda, and Jazz, as a part of their participation in a study. In this particular passage, Cube, Soda, and Jazz try to decide which media they want to use for their presentation and how they want to put the different contents into their presentation.

Cube	Ok ok *Soda* and *Jazz* I think we are stuck here now... 😵 Lets leave that aside for a moment and think first about how we want to present the whole thing. Ok?
Soda	Prezi!
Jazz	Actually there are different possibilities: power point, overhead projector, prezi...
Cube	True you're right *Jazz*... I prefer powerpoint. It's great 🙂
Soda	Powerpoint? But that's really old-fashioned!
Jazz	Yes you're right *Cube*!! @ *Soda*: it's a well-established presentation tool
Soda	I see…
Jazz	Then let's use powerpoint?! How should we set it up?
Cube	Don't know, *Jazz,* but it should really look professional 😎 What would fit with our topic of team-building and the contents of our presentation??
Soda	But why don't we use prezi? I'm sure that it's better and fancier.

Jazz	What do you think about an arrow like structure? ex: solo-players -> method 1 -> method 2 -> method 3 -> ... -> team @ *Soda*: Probably we all can handle powerpoint well.
Cube	Hmmm... I don't know about that, it seems so standard... 🤐 Let's think a bit more about it... maybe we will find an inspiring image for team-building on google...?
Soda	Well then. Google our friend and helper… Id say that the best solution is to think on our own on something metaphorical like from the ego-fighter to the team 😎
Jazz	What about that? img_1_team or that img_2_team ??
Cube	Not bad but I also found some: img_3_team or img_4_team
Jazz	Wow *Cube* image 3 is perfect for us!!!
Soda	I think it's pretty boring. I think it would be better if we created a picture on our own. Then we could also make it fit to my metaphor?!
Cube	I also prefer image 3 *Jazz* but I m not sure yet how we can connect that to all of the team-building methods we are supposed to include in our presentation?! Any other ideas?
Jazz	Yes we could place the problems in an outer circle. In the middle we write team. And in between we put the techniques...or something like that…then the solo-persons metaphorically speaking get together into their own team...how about that?
Cube	That's awesome *Jazz.* Let s do it like that!!
Soda	That's just what I suggested… 😒
Jazz	Great, so glad we figured this out…now we can discuss the details and contents of each main idea in our presentation.... 🙂
Cube	Right *Jazz* 🙂
Soda	Whatever...

Appendix B

Example Stimuli in Study 4. All stimuli were taken from the MR2 database (Strohminger, et al.,

2016). In the essentiality condition, participants were presented with faces that were different

from their own ethnic background (i.e., if the participant was Caucasian, s/he saw Black and

Asian faces). In the non-essentiality condition, the hairstyles of the stimuli persons were

manipulated.

Example Stimuli in the Essential Distinction Condition (Participant Caucasian):
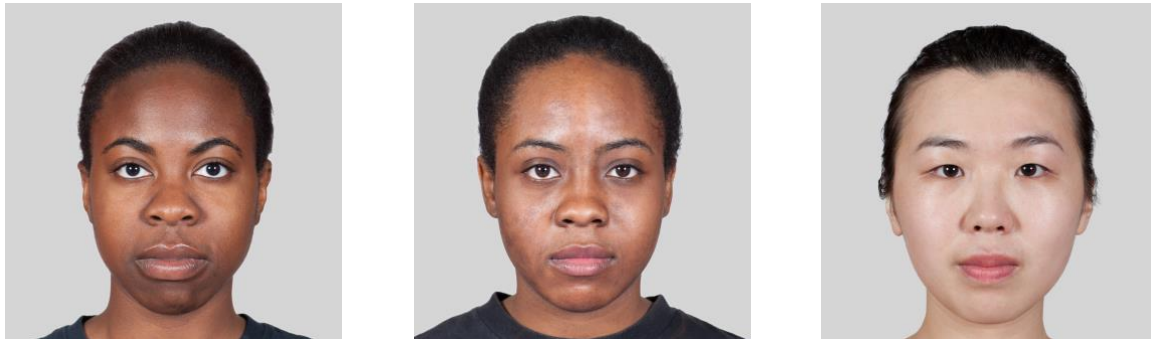
Example Stimuli in the Non-Essential Distinction Condition:

Appendix C

Beginning of the Chat in the norm-violating and norm-consistent target condition in Study 5

Chat Beginning (Norm-violating Target Condition):

Cube	Hey guys, I'm Cube, pleased to meet you!
Jazz	Hey Cube, I'm Jazz. How are you?
Cube	Doing great and you? 🙂
Jazz	Yeah, me too. Looking forward to work on this project with the two of you!
Cube	Same here. I'm sure it will be fun.
Soda	Can we skip the introductions and just get started? I've always hated working in groups.

Chat Beginning (Norm-consistent Target Condition):

Cube	Hey guys, I'm Cube, pleased to meet you!
Jazz	Hey Cube, I'm Jazz. How are you?
Cube	Doing great and you? 🙂
Jazz	Yeah, me too. Looking forward to work on this project with the two of you!
Cube	Same here. I'm sure it will be fun.
Soda	I'm Soda, pleased to meet and work together with both of you too!