Journal of Pacific Rim Psychology

www.cambridge.org/prp

Original Article

Cite this article: Dang J., Liu L., Zhang Q., and Li C. (2019). Leaving an attacked group: Authoritative criticism decreases ingroup favoritism. *Journal of Pacific Rim Psychology*, Volume 13, e7. https://doi.org/10.1017/prp.2019.5

Received: 14 December 2018 Revised: 17 January 2019 Accepted: 18 January 2019

Keywords:

social self-esteem threat; ingroup favoritism; authoritative third party; social identification

Author for correspondence: Caina Li, Email: chinali7371@163.com

Leaving an attacked group: Authoritative criticism decreases ingroup favoritism

Jianning Dang^{1,2} , Li Liu², Qianqian Zhang¹ and Caina Li¹

¹Shaanxi Provincial Key Laboratory of Behavior and Cognitive Neuroscience, School of Psychology, Shaanxi Normal University, Xi'an, Shaanxi Province, China and ²Faculty of Psychology, Beijing Normal University, Beijing, China

Abstract

Research suggests people try to protect their social self-esteem from threats from the ingroup or the outgroup. However, how members react to a threat to social self-esteem from a third party remains unclear. Three studies were conducted to examine the influence of a threat to social self-esteem from an authoritative third party on ingroup favoritism. We explored the effect of negative (versus positive) evaluation from the testing system on explicit and implicit ingroup favoritism in Study 1 and Study 2 respectively. We compared the effect of negative evaluation posed by the testing system or the competitive outgroup on ingroup favoritism in Study 3. Results suggested that individuals experiencing a threat to social self-esteem from an authoritative third party manifested less ingroup favoritism than those experiencing no threat or outgroup threat. The theoretical implications of this research on social identity theory and the practical implications of reducing intergroup bias are discussed.

The recent "spy poisoning case" was an attention-catching event. Some news media speculated that Russia murdered the former spy because he once sold information on dozens of Russian agents to British intelligence agencies (e.g. Fitsanakis, 2018). In addition, in response to London's decision to expel Russian diplomats, Moscow expelled 23 British diplomats and halted the activities of the British Council (Grierson & Wintour, 2018). Social identity theory (SIT; Tajfel & Turner, 1979), especially the self-esteem hypothesis (Hogg & Abrams, 1990; Martiny & Rubin, 2016), provides ideal explanations for these events. Specifically, the deviance of ingroup members and the criticism made by outgroup members may pose a threat to individuals' social self-esteem (Marques, Yzerbyt, & Leyens, 1988; Morier, Bryan, & Kasdin, 2013), and they react defensively to the threat by expressing a tendency favoring their ingroup (Ellemers, Spears, & Doosje, 2002; Hornsey, 2005). Stated differently, group members punish those who betray and pose a threat to their ingroup, and express a preference for their ingroup in response to the threat from the outgroup.

As an authoritative third party, the United Nations has begun to investigate the foregoing spy poisoning case (Pérez-Peña & Rao, 2018). However, the self-esteem hypothesis cannot predict how Russians will react if the investigation reveals that the spy was poisoned by Russia. Inspired by the evidence that group members exhibit less favoritism for their ingroup when the threat to the ingroup's social image is viewed as reasonable (Jost & Burgess, 2000), we propose that the threat to social self-esteem from an authoritative third party may not increase but rather decrease ingroup favoritism. To verify this assumption, we conduct the current research.

Self-esteem and ingroup favoritism

People tend to evaluate the members of the group to which they belong more favorably than members of other groups (Bernstein, Young, & Hugenberg, 2007; Halevy, Bornstein, & Sagiv, 2008; Konrad & Morath, 2012; Yampolsky & Amiot, 2013). This tendency, or ingroup favoritism, can encompass cognitive, attitudinal and behavioral aspects (Hewstone, Rubin, & Willis, 2002). SIT suggests that when social categorization is salient, individuals who strongly identify with their ingroup are motivated to engage in intergroup behavior to maintain or achieve positive distinctiveness of their social groups (i.e. social self-esteem; Tajfel & Turner, 1979). Unlike self-esteem, which refers to overall self-evaluation, social self-esteem is individuals' esteem that derives from their membership in a social group (Crocker & Luhtanen, 1990; Luhtanen & Crocker, 1992; Rubin & Hewstone, 1998). The self-esteem hypothesis (Hogg & Abrams, 1990; Martiny & Rubin, 2016) derived from SIT suggests that threatened social self-esteem motivates ingroup favoritism, as successful ingroup favoritism can provide a positive social image and satisfy ingroup members' need for positive social self-esteem.

Previous research supports the corollary that a threat to social self-esteem increases the tendency favoring the ingroup. On one hand, ingroup members are defensive toward negative

© The Author(s) 2019. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.



evaluations from other ingroup members. Research about the "black sheep effect" suggests that group members reject ingroup criticism and devaluate ingroup critics (Eidelman & Biernat, 2003; Marques & Paez, 1994). On the other hand, negative evaluations from the outgroup result in ingroup favoritism. For example, group members who believe that their ingroup is devaluated by the outgroup allocate more resources for the ingroup (vs. the outgroup) and rate the ingroup (vs. the outgroup) more positively (Hunter et al., 2005). Christians who believed they were evaluated negatively by atheists displayed ingroup favoritism on the ratings of their own behavioral outcomes. Specifically, Christians (vs. atheists) who behaved positively were rated as more positive, and Christian (vs. atheist) members who behaved negatively were rated less negative (Hunter et al., 2004).

Threat to social self-esteem from an authoritative third party

Most available research has examined how people react to a threat to social self-esteem from ingroup members (Hornsey & Esposo, 2009; Morier et al., 2013) or competitive outgroup members (e.g. Hunter et al., 2005, 2004; Hunter, Platow, Bell, Kypri, & Lewis, 1997; Mikulincer & Shaver, 2001). However, both ingroup and outgroup members are actors in the intergroup setting, and intergroup relations often involve more than social actors (Harth & Shnabel, 2015). The influence of negative evaluations from observers (i.e. a third party) on group members' responses has been underexplored. It has been confirmed that actors and observers play different roles in intergroup relations (Seta & Seta, 1992); therefore, we are interested in the effect of a threat to social self-esteem from a third party on ingroup favoritism. Given the essential role a third party plays in decreasing intergroup conflicts (e.g. Fisher, 1983; Harth & Shnabel, 2015), we argue that a threat to social self-esteem from a third party can prevent individuals from being biased towards their ingroup.

Unlike outgroup threat eliciting stronger ingroup identification (Schmid & Muldoon, 2015), previous theoretical perspectives suggest that a threat to social self-esteem from an authoritative third party decreases ingroup identification. Ingroup identification refers to the degree to which people perceive themselves as typical group members, as well as how important and emotionally significant they perceive the ingroup to be to the self (Leach et al., 2008). An authoritative third party is seen as neutral and unbiased (Black & Baumgartner, 1983; Stulberg, 1987). While ingroup or outgroup critics are greeted with skepticism (Morier et al., 2013), a third-party critic is perceived as fairer and trustworthy (Carnevale & Pruitt, 1992). Therefore, the negative evaluation made by an authoritative third party may convince members of the disadvantages of their ingroup. As an ingroup with disadvantages can no longer fulfill its members' need for positive social self-esteem (Tajfel & Turner, 1986), group members tend to leave the group (Bernache-Assollant, Laurin, Bouchet, Bodet, & Lacassagne, 2010; Jetten, Iyer, Tsivrikos, & Young, 2008) or emphasize their heterogeneity with other ingroup members to maintain their individual position in the social system (Ellemers & Barreto, 2001). In sum, the above evidence suggests that when their ingroup is devaluated by an authoritative third party, group members no longer identify strongly with their ingroup (Branscombe, Spears, Ellemers, & Doosje, 2002).

Decreased ingroup identification elicits less preference for the ingroup. Turner and Reynolds (2001) propose that ingroup identification is a necessary and insufficient condition for ingroup

favoritism. Specifically, although high identifiers do not necessarily show ingroup favoritism, low identifiers will definitely not manifest ingroup favoritism. Stated differently, high identifiers can defend their ingroups in many ways, while low identifiers are not willing to defend their ingroup. A series of empirical researches reveals that compared to those who strongly identify with the ingroup, low identifiers manifest less preference for the ingroup (Aberson, Healy, & Romero, 2000; Yampolsky & Amiot, 2013) and more indifference to the ingroup (Ellemers, Spears, & Doosje, 1997). In addition, ingroup favoritism is easier to reduce among low (vs. high) identifiers (Crisp & Beck, 2005).

According to the foregoing theoretical derivations and empirical evidence, a threat to social self-esteem from an authoritative third party decreases ingroup identification and then decreases ingroup favoritism. Therefore, we hypothesize that group members manifest less ingroup favoritism when they receive a negative evaluation of their ingroup from an authoritative third party.

Overview of the present research

Although several studies have investigated how a threat to self-esteem influences ingroup favoritism, the present research is unique in that it sheds light on the source of the threat to social self-esteem rather than the type of self-esteem (Rubin & Hewstone, 1998) or the roles of participants (i.e. intergroup actors or observers; Seta & Seta, 1992). The current research aims to explore the effect of a threat to social self-esteem from an authoritative third party on ingroup favoritism. To examine the hypothesis, three studies were conducted. In Studies 1 and 2, we manipulated a threat to social self-esteem according to a negative evaluation from the testing system (i.e. an authoritative third party) and measured explicit (i.e. trait rating and resource allocation, Study 1) and implicit ingroup favoritism (Study 2). In Study 3, we compared the effect of a threat to social self-esteem from an authoritative third party or a competitive outgroup on ingroup favoritism.

Study 1

In Study 1, we measured attitudinal (i.e. trait rating) and behavioral aspects of ingroup favoritism (i.e. resource allocation). To ensure participants' initial strong ingroup identification, they were asked to select their group membership voluntarily. With reference to a previous study (Mikulincer & Shaver, 2001), a threat to social self-esteem was manipulated according to a negative (vs. positive) evaluation of their ingroup's performance on a cognitive test. Importantly, participants were informed that the evaluation was computed and reported by the testing system (i.e. an authoritative third party). We predicted that a threat to social self-esteem would decrease ingroup favoritism, that is, participants who received a negative evaluation would show less ingroup favoritism than those who received a positive evaluation.

Method

Ethics statement

The study was reviewed and approved by the Committee of Protection of Subjects at the university. Before the study, all participants provided written informed consent and were debriefed at the end of the research according to the established committee guidelines. This procedure was followed in Studies 2 and 3 as well.

Participants

A priori power analysis (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that a sample size of 60 was large enough to provide 80% power to detect an intermediate effect (i.e. Cohen's d=0.65). Seventy individuals were recruited and one participant was omitted because he had previously participated in a similar study. Participants were 69 undergraduates (27 men, 42 women) aged 17–21 years (M=19.29 years, SD=0.88 years) from a large university in China. Participating students received partial course credit. Participants were randomly assigned to either positive or negative evaluation conditions.

Materials

Social categorization manipulation

Following the paradigm used in previous research (Ellemers, Kortekas, & Ouwerkerk, 1999), participants were presented with descriptions of two reasoning styles (inductive or deductive) and asked to select which style best described them. After categorization, participants completed four items (Ellemers et al., 1999) regarding their ingroup identification (e.g. "I feel good about my group"), with responses on a scale from *strongly disagree* (1) to *strongly agree* (7). Higher scores indicated higher ingroup identification. The Cronbach's alpha for this study was .80.

Manipulation of a threat to social self-esteem

Participants were asked to finish a competing task comprising 10 questions that tested whether the induction group or deduction group had the better problem-solving capability. Each question had four options, and participants could choose one option as their final answer. After completing all questions, they received preprogrammed evaluation about their ingroup's performance, which was ostensibly computed and analyzed by the automated assessment system within 5 seconds. In the *negative evaluation condition*, the ingroup scored 57.23 and the outgroup scored 83.67. Given the passing score of 60, this evaluation showed the ingroup possessed poorer problem-solving ability relative to the outgroup. In the *positive evaluation condition*, participants were informed that their ingroup scored 83.67 while the outgroup scored 57.23, with the task passing score being 60; therefore, the ingroup possessed stronger problem-solving skills than the outgroup.

Manipulation check of a threat to social self-esteem

A three-item scale adapted from previous research (Gagnon & Bourhis, 1996; Platow et al., 1997) was used for the manipulation check of a threat to social self-esteem (e.g. "At the moment, I feel good about the problem-solving capability of my group"). All items were answered on a 7-point scale from 1 (*strongly disagree*) to 7 (*strongly agree*). Higher scores indicated higher levels of social self-esteem. The Cronbach's alpha for this study was .83.

Trait rating

Participants were asked to rate their ingroup and the outgroup on six traits (Otten & Wentura, 2001). Participants indicated the extent to which each trait (i.e. friendly, creative, open-minded, reliable, energetic, and empathic) described their ingroup and the outgroup on a 7-point scale from 1 (not at all) to 7 (completely). The order in which participants rated their ingroup or the outgroup first was counterbalanced, and all traits were presented randomly. Evaluations for the ingroup (Cronbach's α = .82) and the outgroup (Cronbach's α = .76) were computed by averaging the ratings on all traits.

Resource allocation

Participants were told that the researcher would spend 1,000 CNY to buy gifts for participants but had no idea how to allocate the money for the two groups; therefore, participants were asked for advice. Participants could choose the amount of money (ranging from 0 to 1,000) for their ingroup, leaving the rest for the outgroup.

Procedure

On arrival, participants were seated in separate cubicles and told that all instructions and tasks would be displayed on the computer and they could only respond with the keyboard. Participants were informed that they would complete a study to investigate how reasoning style influenced problem-solving capability. It was further explained that there were two types of reasoning styles (inductive or deductive), with each reasoning style effective individually, but the two styles might make some difference at the group level. Participants selected their group membership based on the descriptions of the reasoning styles and completed measurement of ingroup identification. Participants were then asked to complete the problem-solving task. At the same time, they were told that their individual performance would determine the performance of their group, so they should answer the questions carefully. After receiving the evaluation (positive or negative), participants completed several measures, including the manipulation check of a threat to social self-esteem, trait rating, and resource allocation. Participants were debriefed and thanked after they finished the experiment.

Results and discussion

Social categorization manipulation check

A t test against the neutral midpoint of 4 showed that all participants strongly identified with their ingroup, M = 5.80, SD = 0.86, t(67) = 17.38, p < .001, at the outset of the study. A Mann-Whitney test¹ revealed the induction, n = 44, M = 5.76, SD = 0.92, and deduction, n = 25, M = 5.86, SD = 0.74, groups identified with their respective ingroup at the same level (p = .518). Therefore, the manipulation of social categorization was successful.

Manipulation check of a threat to social self-esteem

The social self-esteem index underwent a 2 (reasoning style: inductive, deductive) × 2 (evaluation: positive, negative) between-subject analysis of variance (ANOVA). The results showed that neither the main effect of reasoning style (F < 1) nor its interaction with evaluation, F(1, 65) = 1.38, p = .24, was significant; however, the main effect of evaluation was significant, F(1, 65) = 13.39, p < .001, $\eta^2 = 0.17$, 90% CI² [0.05, 0.30]. Participants in the negative evaluation condition, M = 4.05, SD = 0.24, held a lower level of social self-esteem than those in the positive evaluation condition, M = 5.37, SD = 0.26, indicating that the manipulation of a threat to social self-esteem was successful.

Effect of threatened social self-esteem on trait rating

Trait rating was analyzed using a 2 (evaluation: positive, negative) \times 2 (target: ingroup, outgroup) mixed ANOVA, with the second variable as the within-subject variable. The results (Table 1) revealed a main effect of target, F(1, 67) = 28.17, p < .001, $\eta^2 = 0.30$, 90% CI [0.15, 0.42], with the ingroup, M = 4.98, SD = 0.94, rated as more positive than the outgroup, M = 4.16, SD = 0.80. The two-way interaction effect was also significant,

Table 1. Means and standard deviations of trait ratings under different conditions

Trait rating	Negative evaluation condition	Positive evaluation condition
Ingroup rating	4.86 (0.94)	5.12 (0.93)
Outgroup rating	4.36 (0.80)	3.93 (0.76)
Ingroup favoritism	0.50 (1.36)	1.19 (1.27)

F(1, 67) = 4.59, p = .036, $\eta^2 = 0.06$, 90% CI [0.01, 0.16]. Simple effect analysis showed that the ingroup was rated more positively than the outgroup under the positive evaluation condition, F(1, 67) = 25.88, p < .001, $\eta^2 = 0.28$, 90% CI [0.13, 0.41]; however, this tendency was smaller under the negative evaluation condition, F(1, 67) = 5.40, p = .023, $\eta^2 = 0.08$, 90% CI [0.01, 0.19].

Ingroup favoritism on trait rating was computed by subtracting the outgroup rating from the ingroup rating (Yampolsky & Amiot, 2013). A higher score indicated more ingroup favoritism. An independent-sample t test revealed that participants in the negative evaluation condition showed less ingroup favoritism than those in the positive evaluation condition, t(67) = -2.14, p = .036, Cohen's d = -0.52, 95% CI [-1.00, -0.04].

Effect of threatened social self-esteem on resource allocation

A 2 (evaluation: positive, negative) × 2 (target: ingroup, outgroup) mixed ANOVA was conducted, with the second variable as the within-subject variable. This revealed a two-way interaction effect, F(1, 67) = 9.06, p = 0.004, $\eta^2 = 0.12$, 90% CI [0.02, 0.24]. Simple effect analysis (Table 2) showed that the ingroup was allocated more money than the outgroup under the positive evaluation condition, F(1, 67) = 10.01, p = .002, $\eta^2 = 0.13$, 90% CI [0.03, 0.25]; however, this tendency disappeared under the negative evaluation condition, F(1, 67) = 1.04, p = .313, $\eta^2 = 0.02$, 90% CI [0.00, 0.09].

Ingroup favoritism on resource allocation was computed by subtracting the amount of money allocated for the outgroup from that for the ingroup (Platow et al., 1997). A higher score indicated more ingroup favoritism. An independent-sample t test revealed that participants in the negative evaluation condition showed less ingroup favoritism than those in the positive evaluation condition, t(67) = -3.01, p = .004, Cohen's d = -0.73, 95% CI [-1.21, -0.24].

Study 1 revealed that participants under the negative evaluation condition manifested less ingroup favoritism than those under the positive evaluation condition. Stated differently, the results confirmed our hypothesis that a threat to social self-esteem from an authoritative third-party reduced participants' ingroup favoritism.

Study 2

The purpose of Study 2 was twofold. First, we aimed to replicate and extend the findings of Study 1 by measuring the cognitive aspect of ingroup favoritism. Second, to increase the robustness of the findings of Study 1, we measured implicit ingroup favoritism, which was unconscious and less influenced by social norms and social desirability. In this study, ingroup favoritism was operationalized as a shorter reaction time for compatible tasks (ingroup + good/outgroup + bad) than incompatible tasks (ingroup + bad/outgroup + good) in the Implicit Association Test (IAT; Greenwald, McGhee,

Table 2. Means and standard deviations of resource allocation choices under different conditions

Resource allocation	Negative evaluation condition	Positive evaluation condition
Ingroup allocation	487.84 (77.64)	540.66 (66.50)
Outgroup allocation	512.16 (77.64)	459.34 (66.50)
Ingroup favoritism	-24.32 (155.29)	81.31 (133.01)

& Schwartz, 1998). We predicted that participants in the negative evaluation condition would show less implicit ingroup favoritism than those in the positive evaluation condition.

Method

Participants

As indicated in Study 1, a sample size of 24 participants for each condition (positive and negative) was sufficient to achieve 80% power to detect the effect of threatened social self-esteem on ingroup favoritism. However, a priori power analysis may be considered too optimistic (McShane & Bockenholt, 2014), so we recruited as many participants as possible before data analysis. In total, 60 undergraduate students participated in Study 2; 52 participants (25 males) aged 18-24 years, M=19.66 years, SD=0.88 years, were retained after excluding invalid participants with an IAT error rate higher than 20%, as outlined by Greenwald, Nosek, and Banaji (2003).

Materials

Social categorization

Participants were categorized in the same way as Study 1. The 4-item measurement of ingroup identification (Cronbach's $\alpha = .67$) was identical to that used in Study 1.

Manipulation of a threat to social self-esteem

A threat to social self-esteem was manipulated in the same way as Study 1. The three-item manipulation check (Cronbach's α = .84) was identical to that used in Study 1

Measurement of implicit ingroup favoritism

Ingroup favoritism was measured with a computerized IAT (Greenwald et al., 1998), in which participants were asked to categorize words as quickly as possible. The procedure comprised seven blocks: (1) categorizing concept words as relevant to the ingroup (pressing the Q key) or outgroup (pressing the P key); (2) categorizing attributive words as positive (pressing Q) or negative (pressing P); (3) classifying words as relevant to the ingroup or positive (pressing Q) or relevant to the outgroup or negative (pressing P); (4) replicating Block 3, except Block 4 was the official test rather than a practice; (5) categorizing attributive words as positive or negative by pressing switched buttons from Block 2 (i.e. positive = P, negative = Q); (6) classifying words as relevant to the ingroup or negative (pressing Q) or relevant to the outgroup or positive (pressing P); and (7) replicating Block 6 as the formal test. The order of the compatible (Block 4) and incompatible blocks (Block 7) was counterbalanced across participants. The ingroup favoritism index was computed according to previous practice (i.e. $D = lnRT_{incompatible} - lnRT_{compatible}$; Greenwald et al., 2003), with the positive index (D > 0) indicating ingroup favoritism, and the negative index (D < 0) indicating outgroup favoritism.

Procedure

The procedure for Study 2 was almost the same as that for Study 1, except that after receiving an evaluation, participants were asked to complete the word-classifying task (i.e. IAT). Participants were debriefed and thanked after completing the experiment.

Results and discussion

Social categorization manipulation check

A t test against the neutral midpoint of 4 showed that all participants strongly identified with their ingroup, M = 5.52, SD = 0.80, t(51) = 13.79, p < .001. A Mann-Whitney test revealed the induction, n = 36, M = 5.52, SD = 0.76, and deduction, n = 16, M = 5.53, SD = 0.83, groups identified with their respective ingroup at the same level (p = .886). Therefore, the manipulation of social categorization was successful.

Manipulation check of a threat to social self-esteem

The social self-esteem index underwent a 2 (reasoning style: inductive, deductive) × 2 (evaluation: positive, negative) between-subject ANOVA. The results showed no significant main effect of reasoning style or its interaction with evaluation (Fs < 1), but the main effect of evaluation was significant, F(1, 48) = 20.33, p < .001, $\eta^2 = 0.30$, 90% CI [0.13, 0.44]. Participants in the negative evaluation condition, M = 4.05, SD = 1.36, held a lower level of social self-esteem than those in the positive evaluation condition, M = 5.47, SD = 0.68, indicating the manipulation of a threat to social self-esteem was successful.

Effect of threatened social self-esteem on implicit ingroup favoritism

Implicit ingroup favoritism (i.e. D score) was analyzed using an independent-sample t test, with evaluation (2: positive, negative) as the independent variable. Participants under the positive evaluation condition, M = 0.14, SD = 0.18, had higher D scores than those under the negative evaluation condition, M = 0.04, SD = 0.20), t(50) = 1.70, p = .095, Cohen's d = 0.47, 95% CI [-0.08, 1.02]. Furthermore, t tests against zero showed that participants under the positive evaluation condition exhibited implicit favoritism toward the ingroup, t(24) = 3.76, p = .001; however, participants under the negative evaluation condition did not, t(26) = 1.13, p = .269.

Consistent with Study 1, Study 2 revealed that participants who received the negative evaluation showed inhibited ingroup favoritism relative to those under the positive evaluation condition. Importantly, the implicit measurement of ingroup favoritism confirmed that neither social norms nor social desirability could account for our findings. Therefore, Study 2 further supported our hypothesis that a threat to social self-esteem from a third party may reduce ingroup favoritism.

Study 3

In Studies 1 and 2, although the negative evaluation came from the authoritative source (i.e. the testing system), it informed participants of performances of both the ingroup and the outgroup. The manipulation might constitute an intergroup comparison

situation, which possibly made ingroup members experience a threat to social self-esteem from the outgroup. To rule out its possible interference with the results, two measures were performed in Study 3. First, participants in each condition were only informed about the performance of their ingroup. Second, we directly compared the effect of threat to social self-esteem posed by an authoritative third party or the outgroup on ingroup favoritism. We predicted that participants experiencing a threat to social self-esteem from an authoritative third party would manifest less ingroup favoritism than those who experienced a threat to social self-esteem from the competitive outgroup.

Method

Participants

As indicated in Study 1, a sample size of 66 would be sufficient to achieve 90% power to detect the effect of threatened social self-esteem on ingroup favoritism on resource allocation. Eighty participants were recruited. Two participants were excluded because they did not complete all items. The final sample was 78 undergraduates (27 men and 52 women) aged 17–22 years (M=19.31 years, SD=0.94 years) from a large university in China. Participants received partial course credit. Participants were randomly assigned to the third-party evaluation condition (n=39) or the outgroup evaluation condition (n=39).

Materials

Social categorization manipulation

Participants were categorized by cognitive style (field-dependent or field-independent). Following the paradigm used in previous research (Ellemers et al., 1999), participants were presented with the descriptions of two cognitive styles (field-dependent or field-independent) and asked to select which style best described them. After categorization, participants completed two items (Ellemers et al., 1999) regarding their ingroup identification (e.g. "I feel good about my group") on a scale from *strongly disagree* (1) to *strongly agree* (7). Higher scores indicated higher ingroup identification. The Cronbach's alpha for this study was 0.79.

Manipulation of a threat to social self-esteem from different sources

Participants were asked to complete the identical 10-question task as Study 1. After completing all questions, they received a preprogrammed evaluation about their ingroup's performance. In the authoritative third-party evaluation condition, participants received computation results revealing that the ingroup scored 57.23; as the passing score was 60, the ingroup possessed poor problem-solving ability. In the outgroup evaluation condition, participants received results indicating that according to interviews with participants who had completed this test, members of the outgroup generally thought members of the ingroup would have poor performance in the test.

Manipulation check of a threat to social self-esteem from different sources

Previous research indicated that a negative evaluation from an authoritative third party decreased participants' evaluation of their ingroup (Branscombe, Spears, Ellemers, & Doosje, 2002), whereas that from the outgroup decreased perceived evaluation of their ingroup from the outgroup (Hunter et al., 2005, 2004). Therefore, two items developed by Crocker and Luhtanen (1990) were

adapted to use as a manipulation check of the threat to social self-esteem from different sources. The private social self-esteem item ("At this moment, I'm satisfied with my group's problem-solving capability") measured the extent to which members judged their ingroup as being positive. The public social self-esteem item ("At this moment, I think [the outgroup] holds positive attitudes towards [the ingroup] members' problem-solving capability") measured the extent to which members believed that the ingroup was valued by the outgroup. Both items were answered on 7-point scales (1 = strongly disagree, 7 = strongly agree), with higher scores indicating higher levels of social self-esteem.

Resource allocation

The ingroup favoritism on resource allocation was measured in the same way as Study 1.

Procedure

The procedure of Study 3 was almost identical to that of Study 1 except for two changes. First, participants were categorized by cognitive style (i.e. field-dependent or field-independent). Some participants of Studies 1 and 2 reported that they had known something about reasoning style when they were interviewed; therefore, a novel categorization criterion was used to exclude the possible influence of participants' existing experience. Second, a threat to social self-esteem from the authoritative third party or the competitive outgroup was manipulated.

Results and discussion

Social categorization manipulation check

A t test against the neutral midpoint of 4 showed that all participants strongly identified with their ingroup, M = 5.76, SD = 1.03, t(77) = 15.00, p < .001. A Mann-Whitney test revealed that the field-dependence group, n = 53, M = 5.75, SD = 1.09, and the field-independence group, n = 25, M = 5.76, SD = 0.93, identified with their ingroup at the same level (p = .872). Thus, the manipulation of social categorization was effective.

Manipulation check of a threat to social self-esteem

The public self-esteem score was submitted to a 2 (cognitive style: field-dependent, field-independent) \times 2 (source of evaluation: authoritative third party, outgroup) between-subject ANOVA. There was no main effect of cognitive style or its interaction with source of threat (Fs < 1); however, a main effect of source of evaluation was significant, F(1, 74) = 10.24, p = .002, $\eta^2 = 0.12$, 90% CI [0.03, 0.24]. Participants who received a negative evaluation from the outgroup, M = 3.51, SD = 1.47, showed lower public self-esteem than those who received a negative evaluation from the authoritative third party, M = 4.64, SD = 1.75.

The private self-esteem score was submitted to a 2 (cognitive style: field-dependent, field-independent) \times 2 (source of evaluation: authoritative third party, outgroup) between-subject ANOVA. Neither the main effect of cognitive style and its interaction with source of threat was significant (Fs < 1); however, a main effect of source of evaluation was significant, F(1, 74) = 3.98, p = .05, $\eta^2 = 0.05$, 90% CI [0.00, 0.15]. Participants receiving a negative evaluation from the authoritative third party, M = 4.20, SD = 0.95, showed lower private self-esteem than those who received a negative evaluation from the outgroup, M = 4.77, SD = 1.04.

Table 3. Means and standard deviations of resource allocation choices under different conditions

Resource allocation	Third-party evaluation condition	Outgroup evaluation condition
Ingroup allocation	500.00 (60.70)	528.21 (64.68)
Outgroup allocation	500.00 (60.70)	471.79 (64.68)
Ingroup favoritism	0 (121.40)	56.41 (129.36)

In summary, the negative evaluation from the authoritative third party decreased private self-esteem, while the negative evaluation from the outgroup decreased public self-esteem. The results indicated that the manipulation of a threat to social self-esteem threat from different sources was successful.

Effect of threatened social self-esteem on resource allocation

A 2 (source of evaluation: authoritative third party, outgroup) \times 2 (target: ingroup, outgroup) mixed ANOVA was conducted, with the second variable as the within-subject variable. The results (Table 3) revealed a significant two-way interaction, F(1, 76) = 3.94, p = .05 $\eta^2 = 0.05$, 90% CI [0.00, 0.15]. Simple effect analysis suggested that participants under the outgroup evaluation condition allocated more money for their ingroup than the outgroup, F(1, 76) = 7.89 p < .001, $\eta^2 = 0.09$, 90% CI [0.02, 0.21]. Participants under the third-party evaluation condition allocated the same amount of money for the ingroup and outgroup.

The ingroup favoritism on resource allocation was computed by subtracting the amount of money allocated for the outgroup from that for the ingroup. A higher score indicated more ingroup favoritism. The one-way ANOVA revealed that the effect of source of evaluation on ingroup favoritism was significant, F(1, 76) = 3.94, p = .05, $\eta^2 = 0.05$, 90% CI [0.00, 0.15]. Specifically, participants under the third-party evaluation condition manifested less ingroup favoritism than those under the outgroup evaluation condition.

Study 3 provided further evidence for our hypothesis that the threat to social self-esteem from an authoritative third party decreased ingroup favoritism. Specifically, participants experiencing a negative evaluation from the authoritative third party manifested less ingroup favoritism than those experiencing a negative evaluation from the outgroup.

General discussion

The present research examined the effect of a threat to social self-esteem posed by an authoritative third party on ingroup favoritism, using both explicit and implicit measurements of ingroup favoritism. The results of the three studies consistently suggested that the threat to social self-esteem from an authoritative third party decreased people's attitudinal, behavioral and cognitive preference for ingroup members. Specifically, compared to those experiencing no threat or outgroup threat, individuals experiencing a threat to social self-esteem from an authoritative third party manifest less difference in ratings of the ingroup and outgroup, or allocate resources to the ingroup and outgroup more equally. In addition, they showed a weaker cognitive association between the ingroup (relative to outgroup) and favorable attributes.

The present findings challenge the self-esteem hypothesis. The self-esteem hypothesis elaborates the crucial role that self-esteem plays in intergroup bias and suggests that threatened social

self-esteem motivates ingroup favoritism (Hogg & Abram, 1990). However, the current research suggests that a threat to social self-esteem from an authoritative third party does not increase but rather decreases ingroup favoritism. Although our findings are contrary to the self-esteem hypothesis, they complement previous research which suggests that when ingroup standing is perceived as inferior, individuals disidentify with their ingroup (Ellemers & Barreto, 2001; Ellemers et al., 1997). As Turner and Reynold (2001) claimed, the SIT self-esteem hypothesis does not predict the effects of low ingroup status or depressed social identification on ingroup bias. Therefore, individuals' responses to a threat to social self-esteem are more complicated than the self-esteem hypothesis suggests.

The current research contributes to the understanding of the considerable controversy concerning the effect of self-esteem on ingroup favoritism. Some researchers reveal that due to the self-enhancement motive, individuals who are low (vs. high) in self-esteem should be more likely to exhibit intergroup biases (e.g. Branscombe & Wann, 1994; Hunter et al. 2005). However, other researchers find that it is not low but high self-esteem that leads to more in-group favoritism (see Aberson et al., 2000; Hunter et al., 2011). In an attempt to explain this controversy, researchers have argued that enhanced favoritism is likely to emerge when social self-esteem has been experimentally lowered (Hunter et al., 2011), while the pretest level of social self-esteem is positive with ingroup favoritism (Abrams & Hogg, 1988). The findings reported in the current research suggest that it is not enough to take merely the difference of pretest self-esteem or manipulated self-esteem into account, but that the sources of the social self-esteem threat should also be concerned in examining the effect of self-esteem on the intergroup process.

Our findings expand the connotations of SIT by examining how individuals react to the negative evaluation made by an authoritative third party. SIT mainly explains that when experiencing criticism from the ingroup or outgroup, group members try to restore the positive image of their ingroup by devaluing the critic (Eidelman & Biernat, 2003) or displaying ingroup favoritism (Hunter et al., 2005, 2004). However, the current research suggests that such reactions may only occur when the negative evaluation (from the ingroup or outgroup) is regarded skeptically. When the negative evaluation comes from an authoritative third party, the maintenance and restoration of positive group image seems impossible. At this point, individuals' favoritism toward their ingroup will decrease. This mirrors the individual mobility strategy proposed by Niens and Cairns (2003), which points out that individuals are more likely to leave the ingroup with disadvantages when the disadvantages are unchangeable. In summary, how members react to a threat to social self-esteem depends on the source of the threat.

The current research presents practical strategies to control ingroup favoritism and then improve the intergroup relations. Pro-ingroup preference sometimes violates the norm of social fairness. The current research reveals that authoritative criticism can lower ingroup favoritism. Similarly, a recent study (Dang, Liu, Liang, & Ren, 2016) also showed that group members no longer manifested ingroup favoritism when their superior status was criticized as being illegitimate by an authoritative source. It thus can be seen that to suppress individuals' ingroup favoritism and promote social fairness, the role of authoritative third parties is crucial. Following this rationale, the previously mentioned Russian-UK conflict may be resolved if an authoritative third party put its hand up to investigate this case objectively.

The limitations of the current research may inspire future directions. First, use of the minimal group paradigm may reduce the influence of previous experience with membership (Ellemers, Kortekaas, & Ouwerkerk, 1999), but arbitrary categorization and lack of knowledge make membership the only social cue used to guide one's intergroup behavior (Jetten, Spears, & Manstead, 1997). Therefore, future studies should examine the generalizability of our findings in a real situation. Second, we did not directly measure the post-test level of participants' identification with their ingroup, although we believe that decreased ingroup identification can explain why a threat to social self-esteem from an authoritative third party decreases ingroup favoritism. Future research is needed to verify our reasoning. In addition, we employed the feedback provided by the third party and the outgroup to manipulate the source of a threat to social self-esteem. However, these feedbacks may probably be conceived as different in objectiveness. Therefore, alternative methods should be developed and employed to manipulate threat from different sources.

Author ORCIDs. Dianning Dang https://orcid.org/0000-0002-8174-0136

Acknowledgements. We thank Liwen Bianji, Edanz Group China (www.liwenbianji.cn), for editing the English text of a draft of this manuscript.

Financial Support. The authors gratefully acknowledge financial support provided by the Major Project of National Social Science Foundation of China (18ZDA332), the National Social Science Foundation of China (Project Number: 15BRK018), and the Key Project of Philosophy and Social Sciences Research, the Ministry of Education of China (17JZD043).

Notes

- 1. The sample size of deduction group was 25; and the normality assumption of independent samples t test was not met. Therefore, we used a Mann-Whitney test
- 2. As suggested by Steiger (1990), 90% CIs would be more appropriate for η^2 (and R^2). The 90% CI does exclude zero, but barely; a 95% CI would include zero. Furthermore, η^2 (and R^2) cannot be less than zero. Accordingly, Steiger argued that when putting a CI on an ANOVA effect that has been tested with the traditional .05 criterion of significance, that CI should be a 90% CI, not a 95% CI.

References

Aberson C.L., Healy M. and Romero V. (2000). Ingroup bias and self-esteem: A meta-analysis. *Personality & Social Psychology Review*, **4**, 157–173.

Abrams D. and Hogg M.A. (1988). Comments on the motivational status of self-esteem in social identity and intergroup discrimination. *European Journal of Social Psychology*, **18**, 317–334.

Bernache-Assollant I., Laurin R., Bouchet P., Bodet G. and Lacassagne M.-F. (2010). Refining the relationship between ingroup identification and identity management strategies in the sport context: The moderating role of gender and the mediating role of negative mood. *Group Processes & Intergroup Relations*, 13, 639–652.

Bernstein M.J., Young S.G. and Hugenberg K. (2007). The cross-category effect mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science*, **18**, 706–712.

Black D. and Baumgartner M.P. (1983). Toward a theory of the third party. In Boyum K.O. and Mather L. (Eds.), *Empirical theories about courts* (pp. 84–114). New York, NY: Longman.

Branscombe N.R., Spears R., Ellemers N. and Doosje B. (2002). Intragroup and intergroup evaluation effects on group behavior. *Personality and Social Psychology Bulletin*, **28**, 744–753.

Branscombe N.R. and Wann D.L. (1994). Collective self-esteem consequences of outgroup derogation when a valued social identity is on trial. *European Journal of Social Psychology*, 24, 641–657.

Carnevale P.J. and Pruitt D.G. (1992). Negotiation and mediation. Annual Review of Psychology, 43, 531–582.

- Crisp R.J. and Beck S.R. (2005). Reducing intergroup bias: The moderating role of ingroup identification. *Group Processes & Intergroup Relations*, **8**, 173–185
- Crocker J. and Luhtanen R. (1990). Collective self-esteem and ingroup bias. *Journal of Personality and Social Psychology*, **58**, 60–67.
- Dang J., Liu L., Liang Y. and Ren D. (2016). Expected hierarchical integration reduces perceptions of a low status group as less competent than a high status group while maintaining the same level of perception of warmth. Frontiers in Psychology, 7, 2068.
- Eidelman S. and Biernat M. (2003). Derogating black sheep: Individual or group protection? *Journal of Experimental Social Psychology*, 39, 602–609.
- Ellemers N. & Barreto M. (2001). The impact of relative group status: Affective, perceptual and behavioral consequences. In Brown R. and Gaertner S.L. (Eds.), Blackwell Handbook of Social Psychology: Intergroup Processes (pp. 324–343). Oxford: Blackwell.
- Ellemers N., Kortekaas P. and Ouwerkerk J.W. (1999). Self-categorization, commitment to the group and group self-esteem as related but distinct aspects of social identity. European Journal of Social Psychology, 29, 371–389.
- Ellemers N., Spears R. and Doosje B. (1997). Sticking together or falling apart: In-group identification as a psychological determinant of group commitment versus individual mobility. *Journal of Personality & Social Psychology*, 72, 617–626.
- Ellemers N., Spears R. and Doosje B. (2002). Self and social identity. *Annual Review of Psychology*, **53**, 161–186.
- Faul F., Erdfelder E., Lang A.G. and Buchner A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39, 175–191.
- Fisher R.J. (1983). Third party consultation as a method of intergroup conflict resolution: A review of studies. *Journal of Conflict Resolution*, 27, 301–334.
- Fitsanakis J. (2018). Russia tried to kill ex-double spy because he trained Eastern European agencies. *IntelNews*. Available at https://intelnews.org/ 2018/05/14/01%5F2322/, accessed 23 July 2018.
- Gagnon A. and Bourhis R.Y. (1996). Discrimination in the minimal group paradigm: Social identity or self-interest? *Personality and Social Psychology Bulletin*, 22, 1289–1301.
- Greenwald A.G., McGhee D.E. and Schwartz J.L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464–1480.
- Greenwald A.G., Nosek B.A. and Banaji M.R. (2003). Understanding and using the implicit association test: An improved scoring algorithm. *Journal of Personality and Social Psychology*, **85**, 197–216.
- Grierson J. and Wintour P. (2018). Sergei Skripal: Russia expels 23 UK diplomats as row deepens. *The Guardian*. Available at https://www.theguardian.com/world/2018/mar/17/sergei-skripal-russia-expels-23-british-diplomats, accessed 27 July 2018.
- Halevy N., Bornstein G. and Sagiv L. (2008). "In-group love" and "out-group hate" as motives for individual participation in intergroup conflict: A new game paradigm. *Psychological Science*, 19, 405–411.
- Harth N.S. and Shnabel N. (2015). Third-party intervention in intergroup reconciliation: The role of neutrality and common identity with the other conflict party. Group Processes & Intergroup Relations, 18, 676–695.
- Hewstone M., Rubin M. and Willis H. (2002). Intergroup bias. Annual Review of Psychology, 53, 575–604.
- Hogg M.A. and Abrams D. (1990). Social motivation, self-esteem and social identity. In Abrams D. and Hogg M.A. (Eds.), Social Identity Theory: Constructive and Critical Advances (pp. 28–47). London: Harvester Wheatsheaf.
- Hornsey M.J. (2005). Why being right is not enough: Predicting defensiveness in the face of group criticism. European Review of Social Psychology, 16, 301–334
- Hornsey M.J. and Esposo S. (2009). Resistance to group criticism and recommendations for change: Lessons from the intergroup sensitivity effect. Social and Personality Psychology Compass, 3, 275–291.
- Hunter J.A., Cox S.L., O'Brien K., Stringer M., Boyes M., Banks M. and Crawford M. (2005). Threats to group value, domain-specific self-esteem

- and intergroup discrimination amongst minimal and national groups. British Journal of Social Psychology, 44, 329–353.
- Hunter J.A., Kypri K., Stokell N.M., Boyes M., O'Brien K.S. and McMenamin K.E. (2004). Social identity, self-evaluation and in-group bias: The relative importance of particular domains of self-esteem to the in-group. British Journal of Social Psychology, 43, 59–81.
- Hunter J.A., Kypri K., Stokell N.M., Boyes M., O'Brien K.S. and Mcmenamin K.E. (2011). Social identity, self-evaluation and in-group bias: The relative importance of particular domains of self-esteem to the in-group. *British Journal of Social Psychology*, **43**, 59–81.
- Hunter J.A., Platow M.J., Bell L.M., Kypri K. and Lewis C.A. (1997). Intergroup bias and self-evaluation: Domain-specific self-esteem, threats to identity and dimensional importance. *British Journal of Social Psychology*, **36**, 405–426.
- Jetten J., Iyer A., Tsivrikos D. and Young B.M. (2008). When is individual mobility costly? The role of economic and social identity factors. *European Journal of Social Psychology*, 38, 866–879.
- Jetten J., Spears R. and Manstead A.S. (1997). Distinctiveness threat and prototypicality: Combined effects on intergroup discrimination and collective self-esteem. European Journal of Social Psychology, 27, 635–657.
- Jost J.T. and Burgess D. (2000). Attitudinal ambivalence and the conflict between group and system justification motives in low status groups. Personality and Social Psychology Bulletin, 26, 293–305.
- **Konrad K.A. and Morath F.** (2012). Evolutionarily stable in-group favoritism and out-group spite in intergroup conflict. *Journal of Theoretical Biology*, **306**, 61–67.
- Leach C.W., van Zomeren M., Zebel S., Vliek M.L.W., Pennekamp S.F., Doosje B., . . . Spears R. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology*, 95, 144–165.
- Luhtanen R. and Crocker J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. Personality and Social Psychology Bulletin, 18, 302–318.
- Marques J.M. and Paez D. (1994\). The "black sheep effect": Social categorization, rejection of in-group deviates, and perception of group variability. In Strobe W. and Hewstone M. (Eds.), European Review of Social Psychology (vol. 5, pp. 37–68). Chichester, UK: Wiley.
- Marques J.M., Yzerbyt V.Y. and 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**, 1–16.
- Martiny S.E. and Rubin M. (2016). Towards a clearer understanding of social identity theory's self-esteem hypothesis. In McKeown S., Haji R. and Ferguson N. (Eds.), *Understanding Peace and Conflict Through Social Identity Theory: Contemporary Global Perspectives* (pp. 19–32). New York: Springer.
- McShane B.B. and Böckenholt U. (2014). You cannot step into the same river twice: When power analyses are optimistic. *Perspectives on Psychological Science*, 9, 612–625.
- Mikulincer M. and Shaver P.R. (2001). Attachment theory and intergroup bias: Evidence that priming the secure base schema attenuates negative reactions to out-groups. *Journal of Personality and Social Psychology*, **81**, 97–115.
- Morier D., Bryan A.E. and Kasdin L. (2013). The effects of group identity, group choice, and strength of group identification on intergroup sensitivity. *Group Dynamics: Theory, Research, and Practice*, 17, 14–29.
- Niens U. and Cairns E. (2003). Explaining social change and identity management strategies: New directions for future research. *Theory and Psychology*, 13, 489–509.
- Otten S. and Wentura D. (2001). Self-anchoring and in-group favoritism: An individual profiles analysis. *Journal of Experimental Social Psychology*, **37**, 525–532.
- Pérez-Peña R. and Rao P.S. (2018). Putin "likely" ordered Russian ex-spy's poisoning, Britain says. The New York Times. Available at https://www.nytimes.com/2018/03/16/world/europe/uk-russia-spy-putin-johnson.html, accessed 13 August 2018.
- Platow M.J., Harley K., Hunter J.A., Hanning P., Shave R. and O'Connell A. (1997). Interpreting in-group-favoring allocations in the minimal group paradigm. *British Journal of Social Psychology*, 36, 107–117.

- **Rubin M. and Hewstone M.** (1998). Social identity theory's self-esteem hypothesis: A review and some suggestions for clarification. *Personality and Social Psychology Review*, **2**, 40–62.
- Schmid K. and Muldoon O.T. (2015). Perceived threat, social identification, and psychological well-being: The effects of political conflict exposure. *Political Psychology*, 36, 75–92.
- Seta C.E. and Seta J.J. (1992). Observers and participants in an intergroup setting. *Journal of Personality and Social Psychology*, **63**, 629–643.
- Steiger J.H. (1990). Structural model evaluation and modification: An interval estimation approach. Multivariate Behavioral Research, 25, 173–180.
- Stulberg J.B. (1987). Taking Charge: Managing Conflict. Lexington, MA: Heath.

- Tajfel H. and Turner J. (1979). An integrative theory of intergroup conflict. In Worchel S. (Ed.), *The Social Psychology of Intergroup Relations* (pp. 33–47). Monterey, CA: Brooks/Cole Publishing Company.
- **Tajfel H. and Turner J.C.** (1986). The social identity theory of intergroup behavior. *Psychology of Intergroup Relations*, **13**, 7–24.
- Turner J.C. and Reynolds K.J. (2001). The social identity perspective in intergroup relations: Theories, themes and controversies. In Brown R.J. and Gaertner S. (Eds.), Blackwell Handbook of Social Psychology: Intergroup Processes (pp. 133–152). Oxford: Blackwell.
- Yampolsky M.A. and Amiot C.E. (2013). Priming self-determined and non-self-determined group identification: Effects on well-being and ingroup bias. *Group Dynamics: Theory, Research, and Practice*, 17, 137–149.