<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>Social Psychological and Personality Science</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>Draft</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Original Manuscript</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Conflict, Conflict Resolution, Group Processes, Intergroup processes, Intergroup Relations, Political Psychology, Social Identity, Violence, Applied Social Psychology, Behavioral Economics</td>
</tr>
<tr>
<td>Abstract:</td>
<td>Drawing on psychological and economic literature, we investigated how a (high-power) third-party mediator can facilitate the resolution of international conflict through the use of threats of punishment or offers of reward to a conflict party. In the context of U.S. requests to Israel to change its behavior in the Israeli-Palestinian conflict, we found that strongly glorifying Jewish Israelis were most compliant and increased their support for nonviolent conflict resolution when the U.S. threatened punishment to Israel. Meanwhile, weakly glorifying Jewish Israelis if anything were more compliant when the U.S. offered a reward to Israel. The reasons for these diverse reactions and their implications for group processes and intergroup relations, behavioral economics, as well as for conflict resolution through third-party mediation, are discussed.</td>
</tr>
</tbody>
</table>

http://mc.manuscriptcentral.com/spps
Abstract

Drawing on psychological and economic literature, we investigated how a (high-power) third-party mediator can facilitate the resolution of international conflict through the use of threats of punishment or offers of reward to a conflict party. In the context of U.S. requests to Israel to change its behavior in the Israeli-Palestinian conflict, we found that strongly glorifying Jewish Israelis were most compliant and increased their support for nonviolent conflict resolution when the U.S. threatened punishment to Israel. Meanwhile, weakly glorifying Jewish Israelis if anything were more compliant when the U.S. offered a reward to Israel. The reasons for these diverse reactions and their implications for group processes and intergroup relations, behavioral economics, as well as for conflict resolution through third-party mediation, are discussed.

Keywords: intergroup conflict, conflict resolution, mediation, ingroup glorification, punishment.
THE CARROT OR THE STICK IN INTERNATIONAL CONFLICT

Conflict resolution often depends on mediation of negotiations between the conflict parties by a third-party actor. Both scholarly and policy arguments point to potential benefits of outside pressure by third-party mediators (e.g., Barnes & Griffiths, 2008). The South African apartheid regime, for example, only ended when the international community applied external pressure through economic boycotts (Tutu, 2002). Similarly, some third parties have engaged in economic (Barghouti, 2014), diplomatic (Ravid, 2014), and/or academic boycotts (Redden, 2013) of Israel in order to press for a resolution to the Israeli-Palestinian conflict. It has also been suggested that governments, most prominently the U.S. government, should put more pressure on Israel to take more decisive steps towards a lasting resolution to the Israeli-Palestinian conflict, for instance by ending the building of Jewish-Israeli settlements in Palestine (Mearsheimer, 2008).

In line with these arguments and recommendations, recent research found that when a third-party state acts as a mediator, it signals to the disputants that the conflict can be solved, because they assume that the third-party actor would otherwise have refused to mediate (Iwanami, 2014). Despite the apparent importance of third-party mediation for intergroup conflict, there has been no empirical research on what exactly third-party mediators should or can do to facilitate conflict resolution. We thus empirically investigated the effects of different third-party mediating behaviors in the context of American influence on Israel with respect to Israel’s behavior in the Israeli-Palestinian conflict. Specifically, we examined reactions of Jewish Israelis to American threats of punishment or offers of reward if Israel was (not) to change its behavior/policies. Taking a political psychological perspective that citizens can influence their government’s
policies, we focused on Jewish Israelis’ support for, or opposition to, the U.S. request and their own government’s policies.

**Punishment, reward, and cooperation/compliance**

While research on the use of punishment or reward to encourage compliance and cooperation is scarce in the context of intergroup behavior, game theory research has established a wealth of knowledge of the effects of punishment and reward (or prospects thereof) on compliance and cooperation in the context of interpersonal behavior. To investigate these effects, game theory research has developed multiple paradigms based on economic games (e.g. dictator game, ultimatum game). In these games, a participant has to allocate resources between him-/herself on the one hand and “society” in the abstract or concrete other players on the other. This resource allocation is interpreted as an index of selfish vs. cooperative behavior. In different variations of the games, the participant may be punished and/or rewarded by other players for his/her resource allocation.

Despite this research, it remains unclear whether punishment or reward is better to effect compliance/cooperation. Some found that the possibility of punishment increases cooperative behavior (Andreoni, Harbaugh, & Vesterlund, 2003) because the prospect of punishment communicates that a rule is obligatory, whereas the prospect of reward communicates that a rule is voluntary (Mulder, 2008). Others, however, found reward to be equally or even more effective than punishment in promoting cooperation (e.g. Putterman, 2010; Rand, Dreber, Ellingsen, Fudenberg, & Nowak, 2009). One of the reasons for these findings in favor of reward is that the threat of punishment can change...
how people perceive the situation, viewing punishment as the price for self-interested behavior. Then, if the price (i.e. the punishment) is viewed as less costly than cooperation, individuals may choose to act selfishly and accept the punishment as “the price to pay”. If, however, the price is viewed as more costly than cooperation, they will forego selfish in favor of cooperative behavior (Houser, Xiao, McCabe, & Smith, 2008).

Addressing the inconsistent findings in the game theory literature, a meta-analysis based on 187 studies found that reward and punishment can both increase cooperation, yet with a slight advantage for punishment (Balliet, Mulder, & Van Lange, 2011). This advantage, it has been argued, might stem from the fact that in the long term, reward alone cannot sustain cooperation to the same extent that punishment alone can; the threat of punishment continues to influence people, whereas the prospect of reward loses some of its power over time (Sefton, Shupp, & Walker, 2001).

Given the variability in people’s reactions to punishment and reward as reflected in the game theory literature and its focus on interpersonal behavior, it is unclear how the findings of this research map onto real world behavior in the context of international conflict. At the same time, however, it highlights the importance of potential moderating factors that might explain why people do not always respond to punishment or reward in the same way. In the context of intergroup behavior, people’s reactions to the prospect of punishment or reward may depend on their identification with – and particularly their glorification of – the groups to which they belong (e.g. Israel).

**Ingroup glorification as a moderator of reactions to punishment or reward**
The glorification of one’s group, defined as one’s beliefs in the group’s
superiority over other groups and one’s deference to the group’s norms and authorities,
has been identified as an important factor in how people relate to their own and other
groups (Roccas, Sagiv, Schwartz, Halevy, Eidelson, 2008). People who strongly glorify
their group hold, by definition, an inflated image of the group, and they will go to great
lengths to preserve it. If the image is threatened, for example by questionable behavior of
the group or some of its members, high glorifiers will do what they deem best to defend
the group and “save face,” with little regard for the consequences for other groups.
Typically, this defense motivation expresses itself in denial of (Bilali, 2013) or
disengagement from (Leidner, Castano, Zaiser, & Giner-Sorolla, 2010; Roccas, Klar, &
Liviatan, 2006) the (im-)morality of the group’s behavior, the shift from violence-
condemning harm and fairness morals to violence-permitting loyalty and authority morals
(Leidner & Castano, 2012), a reduction in guilt (Roccas et al., 2006) and demands to
redress injustices perpetrated by the group (Leidner et al., 2010). Low glorifiers, on the
other hand, do not defend their group’s questionable behavior and instead are open to
address the group’s flaws and to constructively criticize it (Leidner et al., 2010; Roccas et
al., 2006). Similarly, ingroup attachment, another aspect of ingroup identification
characterized by perceived importance of and commitment to the group, is not related or
even inversely related to defensive reactions to ingroup-committed wrongdoing.

Based on the findings of research on glorification, one could expect that high
glorifiers will respond aggressively to a third-party request threatening the ingroup with
punishment if it does not change its behavior in a conflict with another group. High
glorifiers might find the threat of punishment, signaling that the ingroup’s current
behavior in the conflict is (seen as) morally problematic (cf. Mulder, 2008), more identity- and image-threatening than an offer of reward, and thus be less compliant and more aggressive when exposed to the threat of punishment rather than the prospect of reward. However, situations of third-party mediation are usually characterized by a power differential between the third party and the disputants, with the third party being more powerful. In such situations, we predicted, high glorifiers should view non-compliance and aggression as less viable, because it would be unlikely for the ingroup to be able to avoid punishment. Even if high glorifiers hold an inflated image of their group, this image should be bounded by reality: high glorifying Israelis should still be aware that the U.S. is more powerful. Given the high degree of alignment of high glorifiers’ identity with that of the ingroup, the possible loss of face for the group in the event of punishment should, by extension, also result in a loss of face for the high glorifier him/herself. At the same time, high glorifiers should perceive the threat of punishment from a powerful third party as less severe than a threat of the ingroup’s existence, safety and security posed by other groups (e.g. Palestinians vis-à-vis strongly glorifying Israelis). If they need the third party’s support to survive, they may also ignore any motivation to react aggressively to third-party threats of punishment. For these reasons, compliance should be more face-saving and ingroup-protective than non-compliance.

The prospect of reward, on the other hand, should not be as effective for high glorifiers, as it does not bring the possibility of incurring high, direct costs to the ingroup. In fact, it might even backfire if the reward creates the perception that the ingroup’s behavior is maybe not ideal, but tolerated. Thus, a third-party threat of punishment should be more effective than an offer of reward to induce compliance in high glorifying
members of a conflict party. Unlike high glorifiers, we expected low glorifiers to be more compliant when reward is offered because they would see it as a way to improve their group as well as do what in their eyes is the “right thing”: refrain from violence against and mistreatment of outgroup members.

Method

Participants

We recruited 481 Jewish Israeli adults in Israel through the Midgam project, which has access to representative samples of Jewish Israelis. Eighty-seven participants were excluded because they spent less than 5 minutes on the whole survey (which took on average 25 minutes), and forty-six because they did not pay sufficient attention to the manipulation materials (indicated by manipulation checks), leaving 348 participants for analysis (49% male; relationship status: 33% single, 54% married, 12% divorced, 1% widowed; geographic location: 23% northern Israel, 15% southern Israel, 56% central Israel, 6% West Bank; religiosity: 56% secular, 22% traditional, 13% orthodox, 9% ultraorthodox; occupation: 11% students, 1% soldiers, 67% employed, 13% unemployed, 2% retired, 6% no response; subjective SES: 38% below average, 37% average, 20% above average, 5% no response; 88% Hebrew as native language; 100% citizens of Israel; age: $M = 38.70$, $SD = 13.01$, range = 18-64; religiosity: $M = 5.78$, $SD = 2.10$; conservatism: $M = 4.69$, $SD = 1.57$; seriousness in taking study: $M = 8.52$, $SD = 0.85$).

The exclusion rate of 27.7% of participants is high but not uncommon for online studies, and in fact within range of the exclusion rates that have been deemed necessary to ensure high data quality of online data (Chandler, Mueller, & Paolacci, 2013). The sample size
was determined a priori based on experience with field experiments and online data collection, and the whole sample was collected without any specific stopping rule.

**Procedure**

The experiment was conducted online. Participants were randomly assigned to read one of two vignettes, or to a baseline condition where they proceeded right away to the questions without reading any article. The two vignettes were identical in all respects, except that one depicted the United States offering a reward to Israel to encourage compliance (reward condition), whereas another depicted the United States threatening to punish Israel should it not comply (punishment condition). Participants who read one of the vignettes completed manipulation checks and wrote a brief summary of what they had read about. Participants in all conditions then filled out the measures described below. The procedure and all materials received ethics approval.

**Materials**

*Manipulation checks.* After reading the vignette, participants in the punishment and reward conditions indicated in what manner the U.S. asked Israel to comply (neutral, threatening to punish, offering reward).

*Compliance.* Participants in the punishment and reward conditions responded to eight items tapping to what extent they wanted Israel to comply with the American request (e.g. “Israel should agree to the United States’ demands”; $\alpha = .96$, $M = 4.69$, $SD = 2.21$).

*Support for the ingroup not yielding to other groups’ wishes.* Five items measured how much participants wanted their country to focus on self-interest over other groups’ wishes (e.g. “Israel needs to continue to pursue its own interests regardless of what other
countries want us to do”, “A strong Israel is an Israel that does not bend to the pressure from other countries”; $\alpha = .90, M = 5.89, SD = 1.90$).

Support for the ingroup yielding to other groups’ wishes. Five items assessed how much participants wanted their country to be open to other groups’ wishes (e.g. “Israel needs to be sensitive to the wishes of the world community if it wants to be a true world power”, “Israel will become stronger by listening to its allies”; $\alpha = .87, M = 4.53, SD = 1.64$).

Demands for justice. Five items measured demands to punish the ingroup for transgressions against outgroup members (e.g. “To restore justice, Israel needs to be punished for what it has done”; $\alpha = .92, M = 1.55, SD = 0.89$), adapted from Leidner and colleagues (2013).

Support for coercive-aggressive policies. Four items assessed how much participants supported aggressive-coercive Israeli policies against Palestinians (e.g. “Israel should withhold tax money from the Palestinians if they don't fight terrorism.”; $\alpha = .77, M = 6.83, SD = 2.04$).

Support for military action. Four items assessed participants support for Israeli military action against Palestinians (e.g. “Israel should send military forces to the Palestinian Territories to settle the dispute”; $\alpha = .89, M = 4.48, SD = 2.26$).

Willingness to reconcile. Six items tapped participants’ openness to reconciliation (e.g. “Israelis should try to have friendly relationships with the Palestinians”; $\alpha = .91, M = 5.61, SD = 1.94$).

Ingroup attachment and glorification. Participants’ attachment to and glorification of Israel were assessed with eight statements each, taken from Roccas and...
colleagues (2006). Although attachment ($\alpha = .94, M = 7.80, SD = 1.53$) and glorification ($\alpha = .86, M = 5.92, SD = 1.60$) were intended to be used as moderators, following others (Feygina, Jost, & Goldsmith, 2010; Leidner et al., 2010) we presented these scales after the manipulation so as not to give away the purpose of the study or raise suspicions.

**Results**

Post-hoc power analyses using G*Power (Erdfelder, Faul, & Buchner, 1996) revealed that our statistical power for the interaction effects was .70. While this power was below the recommended standard of .80 (Cohen, 1977), it exceeded the average power in social psychological studies of 0.35 (Bakker, van Dijk, & Wicherts, 2012; Marszalek, Barber, Kohlhart, & Holmes, 2011) and that of .65 for studies published in JPSP, PSPB, and JESP (Fraley & Vazire, 2014).

**Ingroup attachment and ingroup glorification.** Neither glorification nor attachment were significantly affected by condition, $F_s < 0.70, ps > .500$, allowing us to use them as moderators in subsequent analyses, with condition as categorical independent variable (and all possible interaction terms).

The predicted two-way interaction of condition by glorification was significant for all dependent variables. All interaction effects and simple effects are summarized in Table 1 and depicted in Figures 1-7. As expected, high (but not low) glorifiers reported more compliance, support for yielding, demands for punishment of ingroup transgressors, and willingness to reconcile as well as less support for not yielding, support for coercive-aggressive policies, and support for military action when a higher-status third party threatened punishment rather than offering reward. In contrast, low glorifiers reported less support for not yielding, for coercive-aggressive policies and for military action, as
THE CARROT OR THE STICK IN INTERNATIONAL CONFLICT

well as more support for yielding (but not more compliance or willingness to reconcile) in response to a third-party offer of reward. The results among high glorifiers were mostly driven by threat of punishment, not by offer of reward, as the punishment condition also differed from the baseline in three of six cases (while reward never differed from the baseline). The results among low glorifiers, on the other hand, were mostly driven by offer of reward, not by punishment, as the punishment condition never differed from baseline.

Main effects of condition, glorification, and attachment are summarized in Table 2. They were not very relevant for our predictions, but any significant effects were as we would expect based on theory and past research. Two-way interactions of condition by attachment were significant for compliance, support for not yielding, demands for justice, and support for coercive-aggressive policies, such that strongly attached people behaved similar to low glorifiers, whereas weakly attached people if anything behaved similar to high glorifiers (see Supplemental Online Material). While not predicted, past literature has sometimes found attachment to work opposite to glorification (e.g. Roccas et al., 2006). Yet, neither in the literature (e.g. Bilali, 2013; Leidner & Castano, 2012; Leidner et al., 2010; Roccas et al., 2008) nor in the present study are the effects of attachment as consistent as those of glorification, indicating that our effects are first and foremost driven by glorification. The two-way interaction of attachment and glorification, and the three-way interaction of condition by glorification and attachment, were never significant.

Discussion
THE CARROT OR THE STICK IN INTERNATIONAL CONFLICT 12

This study tested the most effective way for third-party mediators to facilitate resolution of international conflict. Strongly glorifying Jewish Israelis were most supportive of nonviolent resolution of the Israeli-Palestinian conflict when the U.S. request was coupled with a threat of punishment, whereas low glorifiers were most supportive when it was coupled with an offer of reward. Consistent with game theory research on the effects of punishment/reward on cooperation, punishment seemed to be slightly more effective than reward (cf. Balliet et al., 2011). These findings extend game theory research from the interpersonal to the intergroup context, and from largely decontextualized economic games to a highly contextualized real-world context. They also highlight the importance of interindividual differences in people’s reactions to punishment/reward, identifying glorification as an important moderator of the effects of punishment/reward on cooperation in the intergroup context.

Extending the social identification, glorification, and intergroup literatures, we further demonstrated that high glorifiers do not always express their propensity to react defensively to threat by aggressive means. Instead, they responded peacefully to both the threat of the adversarial group in the conflict and the threat of punishment delivered by a higher-power third party. This finding showcases that high glorifiers’ defense-motivated intergroup aggression can be avoided, even turned into intergroup prosociality—even under circumstances of extreme threat to security, safety, and identity during conflict. It also suggests that despite high glorifiers having an inflated perception of the ingroup, this perception, while biased, is bounded by reality: in our study, high glorifiers seemed aware that aggression against a higher-power third party is not viable. Last but not least,
our findings have important implications and applications for conflict resolution and third-party mediation.

Of course, our moderation results do raise the important question of what third-party mediators should do to increase the disputants’ willingness to compromise, if some disputants respond better to punishment while others respond better to reward. After all, it is rather unlikely that use of both threat of punishment and offer of reward will be effective, as doing so would send a confusing and contradictory message. If a third-party mediator decides between punishment and reward, punishment may be the better choice, as high glorifiers are a greater obstacle to conflict resolution than low glorifiers (see main effects of glorification in Table 2). Thus, increasing high glorifiers’ support for conflict resolution should take priority. The most effective way to do so is then the threat of punishment, also because it does not seem to backfire on low glorifiers, therefore leading to a ‘net gain’ from the perspective of the mediator. Of course, these recommendations should be taken with a grain of salt, as they stem from only one conflict. At the same time, however, they are based on a field study with great ecological validity and, considering the usual “noise” of field studies, remarkably consistent results. The experimental nature of this field study combined ecological with internal validity. Thus, we believe these results have substantial promise for generalizability.
Acknowledgments

The study reported in this paper constituted the first author’s honors thesis, which was supported by a Research Grant from the Commonwealth Honors College. We would like to thank Amalia Rosenblum for English-Hebrew translations.
References


### Table 1: Two-way interactions of condition and glorification, with simple effects.

Subscripts denote differences between conditions among low and among high glorifiers, respectively; shared subscripts = non-significant, different subscripts = significant.
Table 2: Main effects of condition (Fs and ps) and glorification and attachment (Bs, ts and ps). Subscripts denote differences between conditions; shared subscripts = non-significant, different subscripts = significant.
Figure 1. Compliance as a function of condition and glorification.
Figure 2. Support for not yielding as a function of condition and glorification.
Figure 3. Support for yielding as a function of condition and glorification.
Figure 4. Demands for justice as a function of condition and glorification.
Figure 5. Support for aggressive policies as a function of condition and glorification.
Figure 6. Support for military action as a function of condition and glorification.
Figure 7. Willingness to reconcile as a function of condition and glorification.