Corresponding author

Sabina Čehajić-Clancy, PhD

Associate Professor of Social Psychology

Stockholm University, Department of Psychology

Frescatihagväg 14, 106 91 Stockholm

Phone: +46707899457

Email: sabina.cehajic-clancy@psychology.su.se

Wordcount (main text): 2 915
Moral exemplar intervention: A new paradigm for conflict resolution and intergroup reconciliation

Sabina Čehajić-Clancy, Stockholm University

&

Michal Bilewicz, University of Warsaw

*Current Directions in Psychological Science*

*In press*
Abstract

The challenge of conflict resolution and intergroup reconciliation is maintained by many social and psychological obstacles such as the belief that members of adversary groups are mostly bad and essentially all the same. In this paper, we introduce a novel intervention aimed at challenging these beliefs by exposing people to stories about individuals who have risked some important aspects of their life to save lives of other social groups’ members (so called “moral exemplars”). The effects of this moral exemplar intervention have been tested in several (post)conflict contexts using members of real antagonistic groups with field experiments. Based on these results and in this paper, we discuss three specific and important aspects of the moral exemplar intervention: 1) comparative advantage of the moral exemplar intervention to existing social-psychological interventions aimed at conflict resolution and intergroup reconciliation; 2) content and conditions of the moral exemplar intervention; and 3) implications for future theorizing and research targeting prosocial changes of attitudes and intergroup behavior.

Keywords: conflict resolution, intergroup reconciliation, moral exemplars, prosocial behaviour
During the Nazi German occupation in Poland, providing any help to Jewish people was punished by death. Otto Busse, a devoted member of Nazi party run a contracting office for house painting in Białystok, a Polish city under German occupation. Breaking the occupation rules, Busse payed his Jewish workers the same wages as to non-Jewish ones. After the liquidation of the Jewish ghetto in Białystok, he hid several Jews in his home and adopted a half-Jewish boy, whose father was killed in the ghetto. Later on, he started supporting Jewish underground in Białystok – smuggling German maps, compasses, and pistols. He allowed Jewish partisans to use his office for the production of leaflets and weapons storage (Grossman, 1987). The backbone of research on moral exemplars is that learning about relevant and specific moral exemplars, such as the story of Otto Busse, is an effective approach for changing beliefs and emotions about social groups and in that way improving intergroup relations (Čehajić-Clancy & Bilewicz, 2020).

Human perceptions of social groups as obstacles to successful conflict resolution and intergroup reconciliation

Even after intergroup conflicts have been politically resolved, the challenge of how to build cooperative and trustworthy relations between social groups remains (Halperin & Bar-Tal, 2011). This challenge is further maintained by peoples’ evaluations of social groups. Social-psychological research shows that people tend to evaluate out-groups in a negatively biased and essentialized way by reserving the positive distinctiveness for their own group and perceiving others as mostly bad. Such perceptions of social groups form the basis for negative emotional (e.g., hatred) and behavioral (e.g., discriminatory) responses (e.g., Prentice & Miller, 2007) whilst reducing support for social change, conflict prevention and successful intergroup reconciliation (e.g., Cohen-Chen et al., 2014). Research efforts in social
psychology have produced specific interventions aimed at overcoming these specific barriers including human perceptions and emotions. However, the success of these (and other) interventions is contingent on several important factors such as participants’ attention or lack thereof, their motivation to participate in these interventions and ultimately their motivation to change their initial beliefs (Bar-Tal & Hameiri, 2020). Additionally, the motivation to change is further challenged by psychological tension produced during interventions challenging peoples’ beliefs. The purpose of this article is to introduce a new social-psychological intervention aimed at changing intergroup evaluations through learning about moral exemplars (the moral exemplar intervention). Below we discuss its specific comparative advantage in relation to existing intervention and the implications for future theorizing and research on intergroup relations.

**Social-psychological interventions aimed at conflict resolution and intergroup reconciliation**

In principle, the goal of most social-psychological interventions is to enable prosocial behavioral changes by altering peoples’ attitudes and/or behaviour. Review of existing interventions aimed at attitude changes, in conflict or post-conflict environments, has revealed that such interventions rely on one of the following working principles: actively engaging people with each other (such as contact interventions), exposing people to new information, or instructing people to engage in a specific task (such as perspective-taking interventions). To elucidate this, we have classified key existing social-psychological interventions aimed at conflict resolution or reconciliation according to their basic working principle whilst delineating its specific goal (see Table 1).
Table 1. Social-psychological interventions aimed at conflict resolution and intergroup reconciliation according to their basic working principles. Some interventions employ the strategy of exposing members of different social groups to each other whilst other interventions work on the principle of delivering specific instructions or new information.

<table>
<thead>
<tr>
<th>INTERVENTION</th>
<th>Specific intervention-related goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Principle of active engagement with people</td>
<td></td>
</tr>
<tr>
<td><strong>Intergroup contact</strong> (Pettigrew &amp; Tropp, 2006)</td>
<td>1.1 Bringing people from different social groups into positive and collaborative contact</td>
</tr>
<tr>
<td>2. Principle of instructing people to adopt new perspectives</td>
<td></td>
</tr>
<tr>
<td><strong>Perspective-taking</strong> (Galinsky &amp; Moskowitz, 2000)</td>
<td>2.1 Encouraging people to consider outgroup’s perspective on a specific issue (e.g., conflict)</td>
</tr>
<tr>
<td><strong>Common-ingroup identity</strong> (Dovidio et al., 2014)</td>
<td>2.2 Changing perceptions of social categories from exclusive to more inclusive identification levels</td>
</tr>
<tr>
<td>3. Principle of instructing people to engage in self-reflective exercises</td>
<td></td>
</tr>
<tr>
<td><strong>Self-affirmation</strong> (Čehajić-Clancy et al., 2011)</td>
<td>3.1 Reducing defensiveness in the face of threatening information</td>
</tr>
<tr>
<td>4. Principle of exposing people to new information</td>
<td></td>
</tr>
<tr>
<td><strong>Malleability beliefs</strong> (Halperin et al., 2011)</td>
<td>4.1 Dispelling the idea that groups and conflicts are fixed and unchangeable</td>
</tr>
<tr>
<td><strong>Paradoxical thinking</strong> (Hameiri et al., 2014)</td>
<td>4.2 Inducing realization about the paradox of existing beliefs</td>
</tr>
<tr>
<td><strong>Group-interests intervention</strong> (Gayer et al., 2009)</td>
<td>4.3 Inducing beliefs about ingroup’s losses in the future</td>
</tr>
<tr>
<td><strong>Apology and reparation offer</strong> (Horsney et al., 2013)</td>
<td>4.4 Creating positive intergroup emotional climates</td>
</tr>
</tbody>
</table>

Although the table above indicates discrete intervention goals, in reality most of these interventions propel several important processes. However, and independent of specific working principle or the goal, these and other similar interventions require that people are motivated to either engage, participate, or pay attention to new information or precise
instructions. Furthermore, efficiency of social-psychological interventions is contingent on participants’ motivation to change their beliefs or behavior (Bar-Tal & Hameiri, 2020)

Furthermore, exposure to members of adversary groups (e.g., intergroup contact) or novel information aimed at challenging existing beliefs often produce psychological tension, which can further undermine the success of such interventions. The moral exemplar intervention addresses this particular concern discussed below.

**Development and comparative advantage of the moral exemplar intervention**

Like some of the above listed intervention, the moral exemplar intervention uses the strategy of exposing people to new and unexpected information. More specifically, people are exposed to stories about individuals’ moral and heroic behaviour: members of relevant social groups who have risked some important aspects of their life (e.g., safety, property, family etc.) to save lives of other social groups’ members (moral exemplars). The reason why we focused on stories about groups’ morality is grounded in the research showing that morality is the primary dimension on which individuals and groups are evaluated (Brambilla et al., 2013). Furthermore, in intergroup conflict environments people tend to perceive out-groups (in relation to their own group) as predominantly immoral and evil (Leach et al., 2015). Morality tends to be more attributed to the in-group and often stripped away from out-groups. Such attributions further feed into the vicious cycle of destructive emotions (e.g., hatred) and behavior (e.g., avoidance). At the same time, fear of own immorality is a key obstacle to reconciliation among groups that perpetrated mass crimes (Shnabel et al., 2009) and breeds defensive reactions to reminders about the burden of ingroup’s immoral behavior. Development of the moral exemplar intervention was grounded in this particular line of research pointing to the importance of morality judgments for evaluation of social groups.
The underlying argument was that changing judgements about group morality can be an important way to effectively facilitate prosocial intergroup responses.

In order to provoke inconsistency to existing beliefs about groups’ morality, development of the moral exemplar intervention intended to expose people to unexpected information. This idea was inspired by research in social and emotional learning indicating to violation of expectations (so called “prediction errors”) as a mechanism for driving changes in attitudes and behavior (Hein et al., 2016). Hein and colleagues (2016) show that unexpected prosocial behaviors by an out-group versus in-group individual elicited a relative increase in empathic responses and positive emotions towards new members of the same out-group. Related to this, recent review on social learning demonstrates that such prediction error signals are known to constitute the fundament on which new learning occurs (Olsson et al., 2020) thus inspiring our thinking for new intervention development which would be efficient for changing intergroup beliefs.

Development of the moral exemplar intervention was also inspired by research in education (Han et al., 2017) demonstrating powerful effects of stories as models of facilitating changes in behavior. Finally, and unlike most other interventions relying on the principle of exposing people to new and unexpected information, learning about moral exemplar stories appears to provoke a positive emotional response (Čehajić-Clancy & Bilewicz, 2020; Janković & Čehajić-Clancy, 2021) and thus minimizing the risk of psychological tension usually triggered by violation of expectations. The stories of morally admirable behaviour tend to incite a sense of moral elevation (Haidt, 2000) leading to more pleasant emotional experiences.

Finally, the moral exemplars intervention relies on humans interpreting the social world through constructing narratives. Understanding the meaning and the role of one’s own national or ethnic group is often dependent on existing narrative structures and
personifications (Hilton & Liu, 2017). National metaphors are sometimes confronted by individual life stories and counter-narratives (Bilali, 2013). Therefore, a successful intervention method in contexts affected by conflict could benefit from utilizing a form of a narrative structure. Moral exemplars approach meets this demand of narrative structuring, by presenting stories of individuals who opposed the immorality of group majority.

Drawing on these approaches and existing research evidence, we developed an intervention that operates in the way of providing high inconsistency between the presented information and the held belief (in order to ‘unfreeze’ the beliefs about others) without causing tension (usually produced by the information inconsistency) but rather inducing a positive emotional response (Janković & Čehajić-Clancy, 2021). Induction of such positive affective states produced by violation of expectations might not only sustain participants’ attention but also motivation to learn. In addition, our intervention is easy to implement, has high external validity and can be easily used by practitioners. So far, efficiency of this intervention has been tested with nine experimental studies set in five different (post)conflict societies (Bosnia and Herzegovina, Turkey, Armenia, Poland, Hungary) examining important intergroup reconciliation processes between members of real adversary groups.

In these studies participants were presented with the moral exemplars from national ingroups and outgroups (e.g., Turkish bureaucrat rescuing Armenians during the genocide, German officer offering help to a Polish-Jewish musician, Bosniak hiding a Serbian soldier in his house) utilizing text materials, as well as visual ones (pictures, films). A diverse set of reconciliation-related dependent variables allowed to capture the effects of such exposure on post-conflict intergroup attitudes and behavioral tendencies.

Review of this research shows that learning about moral exemplar stories facilitates several key psychological processes required for successful conflict resolution and intergroup reconciliation (see Figures 1A and 1B; studies 1-2: Čehajić-Clancy & Bilewicz, 2017, studies

**Figure 1A.** The effects of moral exemplar intervention on intergroup attitudes.

*Mean scores with error bars representing standard errors.*
**Figure 1B.** The effects of moral exemplar intervention on intergroup behavioral tendencies.

*Mean scores with error bars representing standard errors.*
Content and conditions of the moral exemplar intervention

Inspired by research in the fields of attitude change (e.g., Deutsch & Fazio, 2008), education (Han et al., 2017) and emotional learning (Hein et al., 2016; Olsson et al., 2020) and based on the published initial evidence (Čehajić-Clancy & Bilewicz, 2020), we propose specific content and criteria for the intervention efficiency. Those are:

1. Usage of **real stories**. Research in education has shown (Han et al., 2017) that using real moral exemplar stories in comparison to conceptual stories is more efficient for promotion of moral emotional and behavior responses. It is essential to note that no artificial materials were presented in any of the intervention studies on moral exemplars: in all cases the interventions were based on factually correct historical information. In the longitudinal studies (Bilewicz & Jaworska, 2013) participants were confronted with living rescuers who shared their stories from the times of the Holocaust.

2. Usage of **relevant stories** set in **relevant socio-political context** focusing on social groups in a meaningful (psychological or structural) relationship. Again, research in the field of education shows that context relevance is the foundation for the effectiveness of such moral exemplar stories. For example, Han and colleagues (2017) demonstrated that stories of relevant individuals (peers) vs. irrelevant individuals (historic figures) promoted greater emulation and change in the behavior. Moral exemplars research in Turko-Armenian context showed that exposure to moral exemplars from another context (Holocaust committed by Nazi Germans) does not affect intergroup attitudes and behavioral tendencies (Witkowska, et al., 2019). Additionally, exposure to moral exemplars does not affect attitudes toward other
unrelated groups (no secondary transfer effects have been observed; Witkowska, et al., 2019).

3. Usage of **extreme counter-information** in order to facilitate arousal and attention required for desired attitude change (Deutsch & Fazio, 2008). Furthermore, in societies where ingroup morality is not threatened, perceptions of ingroup heroism could actually be used as a national alibi for committed crimes and lead to higher prejudice (Bilewicz & Babinska, in press). Therefore, we argue, that exposure to moral exemplars can lead to reconciliation only if the information is a counter-information to the socially shared representation of the history (Hilton & Liu, 2017) or an existing individual belief about group’s morality.

4. Usage of **positive information** to avoid tension and consequent psychological defenses and produce positive emotional responses in the form of moral elevation (Haidt, 2000) and in that way induce or sustain motivation to learn. Furthermore, interventions in which participants were confronted with negative information about ingroup victimization prior to being presented with a film on moral exemplars (Čehajić-Clancy & Bilewicz, 2017) showed more modest effects than other interventions.

5. Presenting outgroup moral exemplars as more **common and typical** to avoid subtyping (Ensari & Miller, 2007). The moral exemplars intervention that presented rescuers in Armenian genocide as extremely marginal proved less effective in eliciting forgiveness and changing perception of current Turks than the same intervention presenting rescuers as being more prototypical (Beneda, et al., 2019).

In conclusion, previous interventions aimed at attitude change through presenting information counter to held beliefs, have often suffered from poor generalizability of the
attitude change due to subtyping, tension provoked by incongruent information or a lack of attention. The moral exemplar intervention overcomes these particular challenges by providing highly unexpected information, which not only evokes attention but also a positive emotional response (e.g., moral elevation), qualities reminiscent of a positive prediction error in formal models of successful learning (Olsson et al., 2020).

Moral exemplars as prosocial regulation of intergroup emotions

When trying to understand the psychological mechanisms at work in this model of reconciliation, we consider the emotion regulation perspective useful for elucidating the appraisal processes underlying learning about moral exemplars. By conceptualizing conflict resolution and intergroup reconciliation as an emotion regulation process (Čehajić-Clancy, Goldenberg, Halperin & Gross, 2016), we argue that the focus of regulation can be on two distinct, yet related, appraisal processes: changing perceptions about social groups (e.g., morality judgements) and changing one’s level of social categorization (e.g., by facilitating perceptions of more inclusive social identities).

Figure 2 proposes such cognitive changes associated with learning about moral exemplars through the theoretical prism of the emotion regulation process whilst acknowledging potentially different mechanisms of learning about ingroup vs. outgroup moral exemplars. We argue that exposure to information which counter existing beliefs about outgroups’ morality might violate expectations about who ‘they’ are whilst at the same time producing a sense of moral elevation associated with prosocial outcomes. On the other hand, information which counter ingroups’ immorality might provide (psychologically needed) affirmation and consequently lead to less destructive and defensive forms of emotion regulation among perpetrator groups (Bilewicz, 2016). Moral emotions such as guilt or
shame – that seem fundamental for reconciliation processes – are effectively downregulated through defensive attributions such as denial or contact avoidance (Leach, Zeinnedine, & Čehajić-Clancy, 2013). We hypothesize that presentation of ingroup moral exemplars, in the context of threatened ingroups’ morality) could constrain the destructive forms of regulating these group-based moral emotions. The focus on ingroup moral diversity could simultaneously tackle the attentional deployment and cognitive change (self-categorization) processes. The narratives of ingroup moral exemplars could create an alternative social identity (self-categorization to a subgroup that is focused around altruistic ingroup prototypes rather than around conflict-facilitating prototypes) and allow to focus on the prosocial aspects of conflict history. Empirical evidence shows that such focus among historical perpetrators restores the sense of moral acceptance and leads to the increase of positive emotions and reconciliatory intentions (Witkowska, et al., 2019). This suggest that the destructive emotion regulation processes among perpetrator groups might be inhibited by such interventions. Social-psychological conditions under which affirmation of ingroup’s morality could lead to more prosocial intergroup outcomes (rather than to moral licensing effects) still remain to be investigated.
**Figure 2.** Theoretical model on prosocial regulation of intergroup relations through learning about moral exemplars: an emotion regulation process. According to this model, prosocial changes at intergroup behaviour level are enabled through creating more prosocial emotions. Learning about moral exemplars can drive important cognitive changes such as changing perceptions about social groups and understandings of social identities, which, in turn, are associated with more positive emotional and behavioural responses.

**Conclusion**

Moral exemplars approach provides a potential alternative to existing models of reconciliation in post-conflict societies. By modifying mental representation of the ingroup or the outgroup, it leads to emotional changes that allow upregulation of prosocial intergroup responses and in that way pave the way to reconciliation between adversarial groups. We argue that the moral exemplars approach is a purely group-based phenomenon. Unlike other approaches to heroism (e.g., Zimbardo, 2011) we do not treat such extreme acts of morality as a trait or a schema that can be primed irrelevant of the group membership or the context. Moral exemplars provide people with important information about the ingroup and outgroup moral variability, which allows transformation of the dominant representation and understanding of the conflict, ultimately restoring positive intergroup relations and establishing ground for successful reconciliation.
Recommended readings

1. Čehajić-Clancy, S., & Bilewicz, M. (2020). (see references) An article presenting the main tenets of moral-exemplars model of reconciliation, as well as presenting several applications of that model in current conflict settings.


Funding

Production of this article has been supported by the Pro-Futura Fellowship by Riksbankens Jubileumsfond (Grant Number: RIK19-1296:3) awarded to Sabina Čehajić-Clancy and Polish National Science Center Grant Sonata Bis awarded to Michal Bilewicz (grant no. UMO-2017/26/E/HS6/00129).
References


