

Science & Society

Improving intergroup relations with meta-perception correction interventions

Samantha L. Moore-Berg^{1,*} and Boaz Hameiri^{1b2}

We explore meta-perceptions (i.e., what we think others think about reality), their impact on intergroup conflict, and the interventions correcting these often-erroneous perceptions. We introduce a two (direct or indirect) by two (with or without framing) framework classifying these interventions, and we critically assess the benefits and constraints of these approaches.

Given the current contentious and polarized climate in many parts of the world, understanding and intervening on the psychological factors that drive conflict is both timely and critical. In this paper, we examine the role of one important psychological driving factor: meta-perceptions. Meta-perceptions are defined as what we think others (in most cases, outgroup members) think about reality. Recent research has shown that meta-perceptions play a critical role in intergroup relations, as they tend to be inaccurate (in fact, overly pessimistic), thus predicting toxic and hostile intergroup relations [1–4]. Based on the tendency of meta-perceptions to be inaccurate, there has been an increasing volume of empirical research assessing interventions that aim to correct these erroneous meta-perceptions, showing their effectiveness in reducing animosity and support for violence in various global contexts [1,5–10].

In this paper, we review this exciting burgeoning literature and propose a new theoretical framework, through which we critically assess these interventions.

Meta-perceptions as a driver of intergroup conflict

Meta-perceptions are often characterized in two different ways. First, they can be what we think outgroup members think about our ingroup – for instance, we might think that the outgroup dehumanizes our ingroup (i.e., meta-dehumanization) or dislikes our ingroup (i.e., meta-prejudice) – or, second, they can be what we think outgroup members think about social issues (also termed ‘false polarization’), such as immigration policy. Due to basic psychological biases and intergroup dynamics, people generally expect that outgroup members will evaluate them negatively [11], which is exacerbated in contexts of conflicting or competitive intergroup relations [2]. This is corroborated by recent research on meta-perceptions in a variety of contexts – people think that the outgroup views them negatively or disagrees with them on social/political issues even if this might not be the case [4].

These overly pessimistic meta-perceptions are highly consequential and have been found to adversely impact intergroup relations, both attitudinally (e.g., showing high levels of reciprocal prejudice and dehumanization toward the outgroup [2]) and in terms of supporting hostile and even violent policies [9]. This is likely the result of people’s basic psychological tendency to reciprocate in kind – they evaluate outgroup members negatively and support hostile actions and policies when they feel that outgroup members evaluate them negatively [1,11]. For example, in the USA, Democrats and Republicans think that the outgroup political party dehumanizes, dislikes, and disagrees with their ingroup about twice as much as in reality. These inaccurate (meta-)perceptions are associated with increased desired social

distancing from the other political party and increased support for policies that limit the rights and freedoms of the other political party [2].

Correcting meta-perceptions

Since meta-perceptions are typically misperceptions, there is room to intervene on them, thereby improving intergroup relations. That is, informing people that they are wrong about what others think may be easier than convincing them that what they think is wrong. After a careful review of the existing literature on meta-perception interventions, we propose a novel two (direct or indirect) by two (with or without framing) theoretical classification of meta-perception correction interventions (Table 1). ‘Direct interventions’ refer to interventions that attempt to correct meta-perceptions intentionally by including factual information that aims to correct exaggerated or false beliefs [1,8,9]. By contrast, ‘indirect interventions’ refer to interventions that have the potential to correct meta-perceptions by providing information about the outgroup’s perceptions through experiential and organic ways. Furthermore, ‘interventions without framing’ refer to interventions that solely include the (direct or indirect) meta-perceptions intervention, whereas ‘interventions with framing’ refer to interventions in which the (direct or indirect) meta-perception interventions are embedded within (typically media-based) messages or narratives. By framing the intervention, researchers emphasize or make salient some elements of the provided information and design the interventions in an engaging and compelling manner, ultimately impacting the interpretation and evaluation of the intervention, increasing engagement with it, and circumventing potential resistance to it [7,10,12].

As an example of a ‘direct intervention without framing,’ one intervention reduced interpartisan dehumanization and animosity

Table 1. The two-by-two theoretical classification of meta-perception correction interventions

		Direct versus indirect	
		'Direct interventions' attempt to correct meta-perceptions intentionally by including factual information.	'Indirect interventions' have the potential to correct meta-perceptions by providing information about the outgroup's perceptions in experiential and organic ways.
With versus without framing	'Interventions without framing' solely include the (direct or indirect) meta-perception intervention.	'Direct interventions without framing' typically first ask participants to estimate how the outgroup perceives reality and then reveals factual information about what the outgroup actually thinks [1,8,9].	'Indirect interventions without framing' typically refer to intergroup contact in which information about the outgroup's perceptions is provided indirectly and in organic ways through dialogue [6].
	'Interventions with framing' embed the (direct or indirect) meta-perception intervention within (typically media-based) messages or narratives to impact the interpretation and evaluation of the interventions.	'Direct interventions with framing' typically embed the factual information into compelling messages or narratives, which are typically delivered using media (e.g., videos) [10,12].	'Indirect interventions with framing' typically refer to parasocial contact interventions, in which individuals can be exposed to what outgroup members think and feel through compelling live or prerecorded audio or video [7].

among Democrats and Republicans by first introducing participants to the concept of meta-dehumanization and then informing them that both partisan groups rate the other side as more human than the other side thinks they do [8]; see also [1,9]). As an example of a 'direct intervention with framing,' researchers found that an intervention that included statistical information (i.e., that more than 90% of Arabs, citizens of Israel, oppose violence against the ingroup members, i.e., Jewish Israelis) embedded in an ostensibly real and compelling newspaper article reduced Jewish Israelis' support of violence against Arabs [10].

'Indirect interventions without framing' most notably refer to intergroup contact. Since meta-perceptions are often inaccurate, providing people with the opportunity to interact with outgroup members and learn about their perspectives may improve intergroup relations by correcting negative meta-perceptions. For example, Americans and Muslims from the Middle East and North Africa who participated in an online contact program that included live encounters and dialogue with the outgroup had significantly reduced meta-dehumanization and dehumanization of the outgroup, even though the program did not directly address these perceptions [6].

'Indirect interventions with framing' typically refer to parasocial contact interventions, in which individuals can be passively exposed to what outgroup members think and feel by watching live or prerecorded audio or video. For example, Colombians who watched a 5-min compelling video featuring FARC (i.e., a leftist insurgent movement that took up arms in 1964 to protect Indigenous and poor communities from exploitation in Colombia) ex-combatants showed significantly fewer misbeliefs about ex-FARC combatants after learning about the ex-FARC's experiences with successful reintegration into society [7].

Strengths and limitations

One main avenue through which correcting meta-perceptions can improve intergroup relations is via changing perceptions. There is a self-reinforcing cycle between meta-perceptions and perceptions – both processes influence each other due to the aforementioned basic psychological tendencies of people to reciprocate in kind [11]. This points to a major limitation of 'direct interventions', namely, that the (largest) effects of these interventions are often found on the most proximal dependent variables (i.e., those that are most closely related to the corrected meta-perception), with considerably smaller or even null

effects on other, more distal, downstream outcome variables [13,14]. For example, researchers found that while their meta-perception correction intervention significantly decreased prejudice, it did not decrease downstream outcomes, such as support for political violence [14]. Notably, when a meta-perception intervention directly corrected perceived outgroup support for violence, it did significantly decrease support for violence against the outgroup as the proximal outcome measure [9,10].

'Indirect interventions (without framing)' that use intergroup contact techniques can be a powerful tool not only to counteract negative meta-perceptions by correcting them but also to improve intergroup relations more broadly [15]. Furthermore, given that they are more holistic and do not correct a specific negative meta-perception(s), we argue that their effect will likely extend to various (proximal and distal) outcome variables. However, as a limitation, negative expectations for intergroup contact (based on, among other things, negative meta-perceptions) can reduce motivation to engage in such interventions or hinder the benefits of intergroup contact, even in instances when the contact is positive, which is not

a guarantee, as it is not fully controlled [5]. Finally, another limitation of both ‘direct’ and ‘indirect’ interventions is that some people might not believe the veracity of the information presented, be it the statistical information presented or the conversation with the outgroup. Meta-perception ‘interventions with framing’ can address these limitations as they provide more control over the interpretation and evaluation of the conveyed messages or narratives in which the (direct or indirect) meta-perception correction interventions are embedded. Thus, the information can be, for example, accompanied with corroborating persuasive evidence or be delivered by a trusted source (such as an ingroup member), and it can be delivered in an engaging and compelling manner, which can also reduce resistance to the intervention [7,10].

One exciting approach that can potentially circumvent the above limitations cleverly combines the ‘direct’ and ‘indirect’ approaches in an intervention with framing. In a 3-min video, entitled ‘Correcting Division Misperceptions,’ which was one of the winning interventions in the Strengthening Democracy Challenge [12], Democrats and Republicans from across the USA reacted, with surprise, to the extent to which they overestimated meta-dehumanization and their disagreement on important social issues. By presenting factual statistical information and by sharing real outgroup members’ reactions, this intervention helped participants learn, directly and indirectly, that conflicting group members share more similarities than they typically think.

Concluding remarks

In conclusion, we hope this paper spurs further interest in meta-perception interventions. Indeed, much work is still needed to understand why, when, and for whom meta-perception interventions are successful. In particular, more research is needed to explore the efficacy of these interventions in different contexts and conflicts, as most of the extant research was conducted in contexts that were largely marked by nonviolent conflict, most notably political polarization [4]. Yet, despite these unanswered questions, existing research points to the promise of meta-perception interventions, which we hope will continue to drive excitement about developing and implementing meta-perception interventions to improve intergroup relations.

Acknowledgments

We are grateful to Emile Bruneau for inspiration; many of the ideas in this article reflect conversations and collaborations with Emile in his effort to put science to work for peace, and we were deeply saddened by his loss to brain cancer on September 30, 2020.

Declaration of interests

The authors have no interests to declare.

¹University of Utah, Salt Lake City, UT 84112, USA

²Tel Aviv University, Tel Aviv-Yafo 6997801, Israel

*Correspondence:
samantha.mooreberg@psych.utah.edu (S.L. Moore-Berg).
<https://doi.org/10.1016/j.tics.2024.01.008>

© 2024 Elsevier Ltd. All rights reserved.

References

1. Lees, J. and Cikara, M. (2020) Inaccurate group meta-perceptions drive negative out-group attributions in competitive contexts. *Nat. Hum. Behav.* 4, 279–286

2. Moore-Berg, S.L. et al. (2020) Exaggerated meta-perceptions predict intergroup hostility between American political partisans. *Proc. Natl. Acad. Sci. U. S. A.* 117, 14864–14872

3. Pasek, M.H. et al. (2022) Biased meta-perceptions about out-partisans’ support for democratic characteristics may erode democratic norms. *Sci. Rep.* 12, 16284

4. Ruggeri, K. et al. (2021) The general fault in our fault lines. *Nat. Hum. Behav.* 5, 1369–1380

5. Borinca, I. et al. (2021) Meta-humanization enhances positive reactions to prosocial cross-group interaction. *Br. J. Soc. Psychol.* 60, 1051–1074

6. Bruneau, E. et al. (2021) Intergroup contact reduces dehumanization and meta-dehumanization: cross-sectional, longitudinal, and quasi experimental evidence from 16 samples in five countries. *Personal. Soc. Psychol. Bull.* 47, 906–920

7. Bruneau, E. et al. (2022) Exposure to a media intervention helps promote support for peace in Colombia. *Nat. Hum. Behav.* 6, 847–857

8. Landry, A.P. et al. (2023) Reducing explicit blatant dehumanization by correcting exaggerated meta-perceptions. *Soc. Psychol. Pers. Sci.* 14, 407–418

9. Mernyk, J.S. et al. (2022) Correcting inaccurate metaperceptions reduces Americans’ support for partisan violence. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2116851119

10. Nir, N. et al. (2023) Kill or be killed: can correcting misperceptions of outgroup hostility deescalate a violent intergroup outbreak? *Eur. J. Soc. Psychol.* 53, 1004–1018

11. Frey, F.E. and Tropp, L.R. (2006) Being seen as individuals versus as group members: extending research on metaperception to intergroup contexts. *Personal. Soc. Psychol. Rev.* 10, 265–280

12. Voelkel, J.G. et al. (2023) *Megastudy identifying effective interventions to strengthen Americans’ democratic attitudes*, Open Science Foundation Published online March 20, 2023. <https://doi.org/10.31219/osf.io/y79u5>

13. Druckman, J.N. (2023) Correcting misperceptions of the other political party does not robustly reduce support for undemocratic practices or partisan violence. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2308938120

14. Voelkel, J.G. et al. (2023) Interventions reducing affective polarization do not necessarily improve anti-democratic attitudes. *Nat. Hum. Behav.* 7, 55–64

15. Paolini, S. et al. (2021) Intergroup contact research in the 21st century: lessons learned and forward progress if we remain open. *J. Soc. Issues* 77, 11–37