11/28/2021

Dr. Reimer,

I sincerely regret that it has taken so long to deal with the serious delay in the handling of your manuscript. I did my utmost to try to have Dr. Adams continue to handle this submission, but it is now clear that this is not possible and thus I have taken over the handling of this manuscript myself. Again, I sincerely apologize for the very lengthy delay and I will do my very best to render a decision as quickly as possible from here.

We have received reviews of the manuscript that you and your co-authors recently submitted to JPSP-IRGP, titled "Meta-Analysis of the 'Ironic' Effects of Intergroup Contact" (MS # PSP-I-2021-0420). All of the reviewers are very knowledgeable and highly respected experts in this general area of research. As will become apparent when you read their critiques, the reviewers have offered several insightful suggestions for ways to strengthen, clarify, and improve your submission. Thus, I am inviting you to revise and resubmit the paper, guided by the suggestions contained in the reviews. I believe that the present paper might be suitable for publication in JPSP-IRGP following some extensive revisions. Because the reviewers' comments are clearly expressed, I will not reiterate all of the points that they have raised.

To submit a revision, go to https://www.editorialmanager.com/irg/ and log in as an author. You will see a menu item called "Submissions Needing Revision". You will find your submission record there.

Make sure to check the appropriate box in the portal to indicate that the paper is a revision rather than a first submission. If possible, I would like to receive your revision within 90 days. If this is not feasible, please email our Peer Review Coordinator, Charlie Retzlaff, at the main editorial office (Cretzlaff@apa.org) with an estimate of when you will resubmit. Longer timeframes are fine.

When you resubmit the paper, please send a detailed cover letter explaining which specific changes you made and which recommendations you did not follow and why. This letter should address all of the points raised in my decision letter plus any other major, non-redundant points mentioned by each reviewer.

In closing, thank you for submitting to JPSP-IRGP. I would also like to thank the reviewers for their service to the field. Their thoughtful comments and suggestions were very helpful in reaching my decision.

I enjoyed reading this paper and I hope you decide to undertake the revision.

Cordially,

Colin Wayne Leach, Ph.D.

Editor, JPSP: IRGP

**Reviewers' Comments**

**Reviewer #1**

I have read the manuscript titled " Meta-Analysis of the 'Ironic' Effects of Intergroup Contact". The ms reports meta-analytic findings from 98 studies with 140 samples. Cumulatively, findings suggest that intergroup contact was associated with less perceived injustice, collective action, and support for reparative policies and these effects change as a function of sample characteristics, measurement type, and type of intergroup relations. Many aspects of this manuscript makes it a strong candidate for publication in JPSP. Analytical strategy is preregistered; large(ish number of studies, focus on moderators, robust methods. However, I would like the authors to address the concerns which I outline below before recommending the ms for publication.

*Theoretical Novelty & Contributions:*

Introduction: Generally well written, but I would have preferred a more direct discussion of the gap and how the paper addresses this gap. More specifically, it would be good to elaborate on how do the findings relate to recent attempts to present cumulative picture of the contact-social change motivations, e.g. Hassler et al 2020. I think it would also be more helpful to amplify on the ironic/sedative effects. A critique of the contact research, e.g. inability bring social change. I like the "why section" and I would like it to be supported with robust evidence. It hinges on a limited number of studies both empirically and theoretically. We need a more elaborate and detailed discussion of why and why not processes. Also, there are other studies that do not fit with any of the main mechanisms through which contact reduces social change motivations, e.g. Cakal et al via perceived threats. A minor point, although the authors might consider it beyond the scope of the present ms, some studies found that in certain contexts contact reduces social change motivations among both advantaged and disadvantaged groups (Cakal et al 2011, Cakal et al 2016).

*Various:*

1. I might be mistaken but the authors use collective action

1. I suspect you forgot a page or line number—I am not sure what your comment refer to.

3. It is not clear which predictor variables are included, quantity, quality, both etc

4. I am a bit sceptical about the cultural distance measure. I can see that it allows the WEIRDness to be incorporated but using US as a referencing point is a generalization too broad. Given that there are too few countries that meets the criteria perhaps the authors would consider dropping it altogether.

1. I will keep it because others found it useful.

2. p.9 I am not sure what the revtools refers to? I can hazard a guess that it is an R package but this info needs to be there

5. Please state which software is used. I know it is R but would be helpful to state this directly.

1. An oversight only someone deep in the analysis could have made.

6. I appreciate the detailed description of the theoretical aspects of the analytical strategy but is it needed? Can it not be included in the appendix?

1. A R3 wants me to expand and R1 wants me to reduce this section, I will keep it as it is unless told otherwise by the editor.

7. I might be mistaken but do the results on three main outcomes include direct, indirect, or both effect types? This is an important discussion which have implications for the hypotheses, moderator analyses, and interpretation of the results.

Discussion: I am not sure if the consensus is still "merging" on the ironic/sedative/paradoxical effects of contact. Again, it would be helpful to clarify whether the effects are direct or indirect. As most studies investigate the indirect effects/associations and this is what the authors allude to when they elaborate on possible mechanisms through which contact reduces social change motivations. If the authors included only direct associations, then this has to be stared (apologies if they did and I missed).

I find the adult vs youth vs children distinction a bit superfluous. To remedy this perception, the authors perhaps can shed more light on specific outcome measures used in these studies and elaborate on their decision to collapse adolescents and children into a single group. They might be similar and within closer age range but I am not aware of any social change attempts by children which is not the case for the youth (see the recent social unrest in Chile and the role of high-school students in kick-starting the demonstrations).

I like the authors recommendation on specifying which types of contact should theoretically affect social change- There is still not enough conceptual clarity with regards to the type of contact.

Last but not least, it would be helpful to have a more sustained and elaborate discussion of how their findings replicate or differ from similar attempts to provide a broader picture of contact-social change motivations path, e.g., Hassleret al 2020.

In sum, I enjoyed reading this manuscript but there are a few issues, detailed above, that need to be addressed before it can be published in JPSP.

Huseyin Cakal

**Reviewer #2**

This is a generally a clear, well-written and rigorous piece of work on a topic that will stimulate broad interest amongst the large community of psychologists - and beyond - who are interested in the relationship between intergroup contact and social change. I enjoyed reading the paper and reflecting on the conclusions drawn by the authors based on their meta-analysis. I think it has considerable promise as a publishable piece of work, though, as will be seen, I also have some misgivings about the argument and conclusions offered.

I'll begin by waiving anonymity as a reviewer: I am John Dixon. Anonymity serves little purpose here as: (a) I have earlier provided both published and unpublished data for this meta-analysis; (b) already know and respect the authors and their work; and (c) have a proverbial 'horse in this race' in that I've contributed to this debate and have my own, albeit developing, position within it. I think it is important to recognize the latter. It inevitably colours my evaluation of how the authors have framed their conclusions. I leave it to the editor's judgement to decide if I have offered tendentious reflections in this respect.

Before proceeding, I should also acknowledge that I am not well-placed to judge the statistical aspects of the meta-analysis. I have never conducted a meta-analysis myself and some aspects (e.g., the use of Bayesian theory and associated ways of presenting key findings), I can only claim to understand in a rudimentary way. I am thus hoping the editor will have engaged other reviewers who have more specific expertise in this regard.

Anyway, on with the review:

Why I liked the paper

\* The paper is timely and well-conceived. It is timely because the idea of the 'ironic effects' of contact on the political attitudes of the disadvantaged has been generating growing interest over the past decade or so and has thus also been generating a burgeoning empirical literature. However, as the authors point out, that literature is currently producing more complexity, dissensus and division than clarity or integration. There is also as much heat as light. Work of synthesis is much needed and that is where this paper contributes. The paper is well-conceived in that the authors generally do a good job in distilling the main themes and unanswered questions of this emerging literature. They also do a good job in tracking down relevant empirical work, both published and unpublished. It is a fast-evolving field with lots of emerging papers.

\* The written quality of the paper is consistently good and clear. The attention to methodological and technical detail is patient and careful.

\* The paper clearly results from an enormous amount of work conducted in a short space of time. I know that this is not per se a reason to publish. However, I find hard work conducted at pace admirable, especially when dealing such as with a fast-moving literature in order to construct an up-to-date, timely review.

\* Can I also add that the map and associated data presented as Figure 2 is very revealing. It captures an awful lot in terms of what we (think we) know about intergroup contact relations on this particular topic, e.g. where it has been studied and using what methods (e.g. I found the almost complete absence of experimental work intriguing). Basically, it seems we know quite a lot about relations in the US, a bit about some parts of Europe and some parts of the Middle East; a tiny bit about India (bearing in mind the size and diversity of India in terms of intergroup relations); and not much about anywhere else, including most of Africa, China, Russia and South America, i.e. most of the world. To say the least, this map in itself warrants the authors' conclusion that we need to know more about "…the direction and magnitude of the associations between intergroup contact and support for social change across cultures, groups, and issues." Is it worth flagging this graph again in the conclusion/discussion section as I think this point warrants some further elaboration?

Some thoughts on the classification and nuance with respect to predictor, outcome and moderation variables

\* I'll admit I had some misgivings about how some of the key variables that organized the meta-analysis were classified, notably the key predictor of contact. The findings of the moderation analysis showed that direct contact had the strongest effects. I'm not surprised to be honest. The other proxy measures of contact, in this context, were not convincing to me, i.e. mere opportunity for contact and number of friends not from the ingroup (with the possibly strained assumption that this is systematically related to the amount of positive contact experienced with specific historically advantaged outgroup members). Both of these variables operationalise the kind of contact at the heart of this meta-analysis in a way that is subject to numerous challenges, i.e. opportunity for contact does not necessarily translate into actual contact and, even if it does, how do we know this contact is either positive or frequent? As such, I would like to see what the overall results of the meta-analysis look like without these proxy measures.

\* I would also like to see what the results look like when using clear measures of direct, qualitatively positive contact as the main predictor which, in my opinion, fits better with the logic of the 'irony of contact' argument (as well as with the main explanations of the effect reviewed in the paper's introduction). In any research I've done, qualitatively positive contact is far more predictive of the 'irony of contact' effect than mere quantity of contact measures (which often leave contact valence a bit vague). Again, this distinction seemed to be collapsed in the main meta-analytic results, unless I've missed something?

In short, what happens to the meta-analytic trends when we test the strong case for ironic effects of direct, qualitatively positive contact between adults, without including indirect proxies of such contact, and how does this vary across key contexts (e.g. colonial versus other environments)?

\* Finally, I found the division of different forms on unequal relationship into categories a bit opague - notwithstanding the reliability of coding checks - and I would have liked richer exemplification of this process as it is high significant in terms of interpreting the results. Some of the categories are not as independent as they are presented in my opinion. For example, many so-called post-colonial relations are also implicated in relations of short term and long term migrancy. In places like the UK, to use one example, the colonized effectively migrated to the centre - and many members of Afro Caribbean, Pakistani, Bangladeshi, and Indian communities are thus both migrants and historically colonial subjects.

\* Can I add an additional point about framing interpretations here. I felt the authors sometimes implied, albeit subtly, that the fact ironic contact effects were strongest amongst colonial, post-colonial or recent immigrant communities somehow diminished the importance of this tradition of work and its key findings. I would argue quite the opposite. These are the communities that have suffered, and continue to suffer, from the most extreme forms of inequality (e.g., Palestinians right now) and thus also communities where motivations to enact political struggle are of paramount importance.

Some thoughts on the paper's discussion and framing

The authors argue that there are convincing alternative explanations of the irony of contact effect and imply, indeed, these are equally plausible at various points. They highlight in particular that this effect wanes or disappears when negative contact experiences are controlled for. I accept that this is an interesting point, but I have reservations about how it is currently framed, and I disagree that this is an equally plausible explanation at this stage. First, the argument offered in the paper is basically statistical. That is, the authors do not offer much in the way of providing compelling conceptual and analytic grounds for the role of negative contact in this context. Second, the applied implications of this alternative explanation are unclear to me. These are important to clarify further because the contact literature is essentially a concerted attempt to reduce prejudice and increase social equality. For example, if the 'irony of contact' effect diminishes or disappears when disadvantaged group members also experience negative contact what does this mean in practice in terms of devising contact interventions in contexts such as schools or workplaces? Should we only encourage positive contact in situations where it is offset by negative interactions that encourage political activism and pre-empt such ironic effects? I doubt that is what the authors would argue, but then they need to be clearer about their own position here. Third, although there is some statistical evidence supporting the role of negative contact in offsetting the ironic effects of contact, which the authors present, it is worth noting that: (a) the relevant literature on negative contact here is tiny and largely undeveloped as an explanation of collective action, perceived injustice or policy support, and (b) the wider role of negative and positive contact interactions in shaping other important indicators (e.g. prejudice) is still hotly debated and thinly evidenced. Some commentators, for example, claim that positive early experiences of intergroup contact offset the effects of negative contact on prejudice in later life (see Schäfer et al.'s (2021) review of research on negative intergroup contact experiences published in the JSI). In sum, I think the authors need to acknowledge the complexities more carefully here and be more cautious about drawing their conclusions.

In terms of drawing out theoretical interpretations and implications, I felt that two further points could be considered. First, although I appreciate the value of a tight focus, contact experiences are invariably relational and evidence on the responses of disadvantaged groups require some consideration of relational dynamics with advantaged. For example, there is an interesting body of work (e.g. Julia Becker, Steve Wright) that suggests that positive contact that, in terms of communication content, involves recognition of illegitimate inequalities and group differences may not lead to ironic effects on disadvantaged group activism and, moreover, may also produce allyship amongst the historically advantaged. Second, I want to highlight that an interesting and large-scale piece of work has just been published in the JPSP that offers a 'needs based' reconciliation of some of the tensions explored in the present paper (Tabea Hassler and colleagues). This is a follow up to their earlier paper in Nature Human Behaviour, which is cited in the present paper and also offers a relational account in that it identifies a specific form of contact that encourages both members of historically advantaged and disadvantaged groups to work towards social change to combat inequality.

Finally, and again as a matter of providing a more nuance contextualization of the metanalysis, it is worth noting that many of the key conceptual papers cited in this article did not merely offer a critique of the contact hypothesis. They also offered broader reflections on the limitations of the prejudice reduction model of social change on which it is based, which presupposes that inequality and discrimination decline mainly through fostering warmer feelings and thoughts amongst the advantaged towards the disadvantaged. The ironic or paradoxical effects of prejudice reduction interventions need to be framed as part of this wider debate, which includes, for example, work on intergroup helping, common identification, benevolent sexism, positive stereotyping, work on moral emotions such as hope, and so on. I recognize fully one cannot include everything in a single paper, but this point is worth at least a footnote, especially given that the work of Reicher, Wright and others is flagged so directly in the paper's introduction.

A minor point

\* The authors used the term 'observational' throughout the paper to classify studies that I would treat as 'surveys' or 'questionnaire' studies. Maybe this is a formal technical term in meta-analyses these days? I tend to treat observational studies as involving direct observations of behaviour rather than self-report measures and that why I found the terminology confusing in the beginning. I worked it out.

Conclusion

Overall, I want to emphasize that my points above are offered in a spirit of constructive engagement and with due recognition that this is a strong and timely meta-analysis. I very much enjoyed reading it and feel that is generally a rigorous piece of work that already makes a contribution to the field. I wish the authors the very best of luck in developing their work.

John Dixon

**Reviewer #3**

The manuscript 'Meta-Analysis of the 'Ironic' Effects of Intergroup Contact' is a systematic analysis of the association between i) contact and collective action, ii) contact and support for redistributive policies, and iii) contact and perceived injustice. The authors used a large dataset of both published and unpublished data to test for the robustness of their results.

I congratulate the authors for collecting such a large dataset and providing a detailed report. Further, I appreciate that the authors preregistered their analyses and uploaded their materials on OSF. Overall, the manuscript summarizes a growing body of evidence. Yet, I have some concerns that lower my enthusiasm to publish this paper in its current form. Below, I will elaborate on each of my concerns.

Preregistration and Data Inclusion

- The authors list their deviations from their preregistration only in the SM and not in the main manuscript. Please list major deviations (e.g., the authors registered to run 6 [without moderator] models but reported only 3) in the main manuscript as well

- Almost half of the data (n <= 100.000) are collected one by of the authors [the names are publicly available] and not (yet) published. I consider this problematic, particularly given that unpublished (compared to published) data is given more credibility (e.g., p.9 - discussion on publication bias). Please clearly state this in the main manuscript. Furthermore, please provide information on how and why (of the ongoing meta-analyses or other reasons) this data has been collected. Ideally, the corresponding datasets (Sengupta, 2020a; Sengupta, 2020b) should be made available as well.

- The authors stated that they would update their data set every 4 months until submitting their manuscript. However, it seems that at least two large studies (advance online publication in 2020) are missing that should have already been available:

o Effect of contact on policy support: Earle, M., Hoffarth, M. R., Prusaczyk, E., MacInnis, C., & Hodson, G. (2021). A multilevel analysis of LGBT (Lesbian, Gay, Bisexual, Transgender) rights support across 77 countries: The role of contact and country laws. British Journal of Social Psychology, 60(3), 851-869.

o Effect of contact on perceived illegitimacy: Hässler, T., Ullrich, J., Sebben, S., Shnabel, N., & Bernardino, M. (2020). Needs satisfaction in intergroup contact: A multi-national study of pathways toward. Behavioral and Brain Sciences, 35(6), 411-425.

Please include these studies in the analyses.

Introduction

- The authors do not seem to adequately represent the current evidence regarding the positive and negative effects of contact on social change as the main take away is that negative and positive contact effect seems equally likely.

"Supporting this argument, initial studies provided evidence that contact with advantaged group members reduces disadvantaged-group members' perceptions of injustice (Saguy et al.,2009), collective action (Çakal et al., 2011), and support for reparative policies (Dixon et al., 2007). Others, however, have pointed to conflicting evidence (Poore et al., 2002) and argued that intergroup contact might instead make discrimination more salient and, thereby, increase disadvantaged-group members' support for social change (Pettigrew et al., 2011)", p.2.

"As societies are becoming more diverse, there are more opportunities for members of different groups to come in contact with each other. Psychologists disagree about what this

means for disadvantaged-group members' awareness of and opposition to social injustice. On the one hand, researchers have argued that intergroup contact might reduce disadvantaged group members' perceived injustice, discourage them from engaging in collective action, and diminish their support for reparative policies. On the other hand, researchers have argued that intergroup contact might not reduce, or even increase, support for social change in disadvantaged groups. Settling this debate is important both because intergroup contact is a promising means of reducing prejudice (Pettigrew & Tropp, 2006) and because political mobilization by the disadvantaged is often crucial for achieving social change (see, for example, Mazumder, 2018)." p. 6.

First, the authors contradict themselves in their discussion "There is an emerging consensus that intergroup contact has the 'ironic' effect of undermining support for social change in disadvantaged groups. …. Thus, our meta-analysis seems to support the emerging consensus", p.27. Second, many studies (studies included in the current meta-analyses and the study by Earle et al., 2021 that the authors missed) show evidence for a significant but small negative association between contact and collective action/policy support, which has also been discussed in several published review papers (three of them are cited by the authors). I would recommend that the authors represent published work more adequately and tone down their novelty claim.

*Analysis Strategy*

I am surprised that the authors centered the prior to 0. Also, why would 50% of the cases lay outside of the interval? Based on the existing evidence (trend toward irony of harmony effect) and the negative effect reported in Pettigrew and Tropp's (2006) meta-analyses, I would have expected a negative effect. While the authors preregistered their decision, I am curious to know whether the results are robust across various priors.

1. It is good practice (for a longer argument, see Gelman et al., 2017) to use prior distributions that are *conservative* (i.e., they do not bias results in either a positive or negative direction) and *weakly informative* (i.e., they incorporate *some* domain knowledge to regularize inferences) in a given domain. The Normal(0, 0.31605) distribution is conservative because it is centered around 0 and is weakly informative because, even though it is centered around the *a priori* most plausible values, implausible effect sizes still have some prior probability. For example, an *average* effect size of | *r*mean | > .40 still has a prior probability of Pr(*r*mean > .40) = 18% in our analysis. In their insightful classification of effect sizes, Funder and Ozer (2019) considered an effect size of *r* > .40 “a very large effect size in the context of psychological research” and “likely to be a gross overestimate that will rarely be found in a large sample”. In any case, our results are robust to choosing substantially more and less informative (i.e., wider) prior distributions, as reported in the Robustness Checks subsection. Further details on how we chose the prior distributions is available in the preregistered protocol.

- The current introduction to Bayes (p.16) is not sufficient to understand Bayesian statistics. I would recommend to either expand this section or referring to relevant introductory articles.

Results

- Meta-Bias (p. 24): It would be helpful if the authors provide some details about the different estimates (i.e., RMA, PET-PEESE, 3PSM, p-uniform) they use so that the reader can understand in which way these estimates differ.

1. I will add *some* explanation but will leave the rest to the reader.

- Alternative predictors (p.25): The authors assessed the association between residualized positive and social change and residualized negative contact and social change. Because the effect sizes are larger for negative contact, they conclude that they find stronger evidence for negative contact. Given that previous work has discussed the possibility of a contact threshold (e.g., MacInnis & Hodson, 2019) and interaction effects (Árnadóttiret al., 2018), I suggest that the authors also test for interaction effects.

*Discussion*

- Smaller effect sizes for collective action vs. prejudice. While the former is a behavioral (intention), the latter is an affective/cognitive construct. Could that also partly explain the relatively small effects? I would also expect that longitudinal studies (where you parceled out the autoregressive effects) had smaller effect sizes compared to cross-sectional studies.

1. It might be one reason, yes. But, as we note in the paper, it doesn’t make much sense to compare effect sizes across outcomes.

- Several moderators (e.g., supporting contact, discussion of inequalities, needs satisfaction, dual identity) of the irony of harmony effect have been empirically tested or theoretically discussed. The authors should refer to these alternative predictors in their discussion

1. We made the conscious choice to limit the scope of our analysis to the direct relationships between intergroup contact and support for social change in disadvantaged groups.
2. by not considering within-study moderators or mediators.

Minor comment

- The 'sexuality' group should be renamed in the main manuscript and the SM. The samples include both sexual and gender minorities. Importantly, sexual orientation and gender identity are distinct constructs. E.g., a trans man can be heterosexual. Thus, this person would be a gender minority (because he is not cisgender) but not a sexual minority member (because he is heterosexual). An alternative label could be LGBTIQ+ or sexual and gender minorities.