Protocol:
Meta-analysis of the
'ironic' effects of intergroup contact

March 22, 2019

Abstract

This is the preregistered protocol for our meta-analysis of the relationship between intergroup contact and support for social change among members of disadvantaged groups. This protocol follows the PRISMA-P guidelines (Moher et al., 2015; Shamseer et al., 2015) for preparing a systematic review protocol.

1   Introduction

1.1   Rationale

In this meta-analysis, we are considering more than a decade of empirical research (Dixon, Durrheim & Tredoux, 2007) on the potential 'ironic' (Saguy, Tausch, Dovidio & Pratto, 2009), 'sedative' (Cakal, Hewstone, Schwär & Heath, 2011), or 'demobilising' (Dovidio, Love, Schellhaas & Hewstone, 2017) effects of intergroup contact.

Prior research mostly studied intergroup contact as a means to reduce prejudice and social conflict (for a meta-analysis, see Pettigrew & Tropp, 2006). Most research on intergroup contact had studied members of advantaged or 'majority' groups (for a meta-analysis, see Tropp & Pettigrew, 2005). Reducing prejudice among advantaged-group members had been advocated as a way to reduce discrimination and to foster positive social change. Following initial theorising by Wright (2001), Dixon, Durrheim and Tredoux (2005), and Reicher (2007), researchers begun studying whether, in historically unequal societies, intergroup contact could instead *hinder* social change by reducing support for social change among members of disadvantaged or 'minority' groups. Since Dixon et al.'s (2007) first empirical test of this hypothesis, a growing number of studies have examined the relationships between intergroup contact and social change among disadvantaged-group members.

To date, this literature has been the subject of narrative reviews (for example, McKeown & Dixon, 2017), but not of a systematic review or quantitative synthesis. Filling this gap, this protocol outlines a systematic review and meta-analysis of the so-called 'ironic' effects of intergroup contact. This endeavour allows us to quantify the strength of the relevant relationships and to identify potential shortcomings of the available evidence.

1.2 Objectives

In this meta-analysis, we examine the relevant literature to answer the following research questions:

Q0 *Does intergroup contact help or hinder social change?* For this review, we define social change as a reduction in social inequality that improves the situation of disadvantaged groups in a society. Though this question motivates our review, the meta-analysis will not conclusively answer whether, on balance, intergroup contact helps or hinders social change. By answering Q1–Q3, however, our review will test the most prominent hypotheses explaining how intergroup contact *could* hinder social change.

Q1 *Does contact with advantaged-group members reduce collective action among disadvantaged-group members?* Our review uses Wright's (2010) definition of collective action as any action taken to improve the ingroup's position in society. As such, collective action (by disadvantaged-group members) challenges the status quo and, if successful, improves the situation of the disadvantaged ingroup. To the extent that it prevents collective action, intergroup contact hinders social change.

Q2 *Does contact with advantaged-group members reduce disadvantaged-group members' support for policies that would benefit their ingroup?* Redistributive policies (for example, affirmative action) and similar initiatives can be important means to reduce social inequality. Members of disadvantaged groups, who stand to benefit from their implementation, are important advocates for these policies. To the extent that it reduces support for such policies, intergroup contact hinders social change.

Q3 *Does contact with advantaged-group members reduce perceptions of injustice among disadvantaged-group members?* Some researchers (e.g., Saguy et al., 2009) suggest that intergroup contact could lull disadvantaged-group members into a false sense of equality. Being aware of the injustice faced by the ingroup, however, is an important prerequisite to instigating social change (van Zomeren, Postmes & Spears, 2008). In addition, accurately perceiving the social reality is likely important to coping in a society marked by inequality. To the extent that it reduces perceptions of injustice, intergroup contact hinders social change.

For Q1–Q3, we explore the following questions:

QA *What is the direction of the relationship between intergroup contact and the outcome variable?*

QB *How strong is the relationship between intergroup contact and the outcome variable?*

QC *How much does the relationship vary across study designs, measures, and settings?*

Beyond the meta-analysis, we use the information collected on each study to review the quality of the available evidence and to identify gaps and limitations in the literature. Below, we describe the methods used to find, describe, and synthesise relevant studies.

2 Methods

2.1 Eligibility criteria

In this meta-analysis, we are considering studies that *i.* included **participants** of a relatively disadvantaged group, see *Types of participants*, that *ii.* manipulated or measured **intergroup contact** with a relatively advantaged group, see *Types of predictor variables*, and that *iii.* measured **one or more** of: **collective action** (intentions), **support for policies** that benefit or harm the participants' ingroup, and/or perceptions or expectations of **injustice**, see *Types of outcome variables*. Below, we discuss these criteria in detail.

2.1.1 Types of studies

We include quantitative studies of all designs, as long as they fulfil the criteria laid out above. We consider study design features as possible moderators of study effect size.

2.1.2 Types of participants

We include samples of participants whose ingroup is disadvantaged (in terms of status, power, and/or resources) relative to the outgroup they have (or report to have) contact with.

2.1.3 Types of predictor variables

We include studies if they measure or manipulate the quantity of, quality of, and opportunity for contact with members of outgroups that are relatively advantaged (in terms of status, power, and/or resources) compared to the participants' ingroup (see *Types of participants*). We consider both positive and negative contact, where both are measured. We consider opportunity for contact, as it is a potential precursor to face-to-face contact, but not imagined or indirect contact.

2.1.4 Types of outcome variables

We include studies that measure **at least one** of the following outcome variables:

Collection action. We consider measures of observed, reported, or intended engagement in any action aimed at improving the position of the participants' ingroup in society, including participating in protests, signing petitions, and other forms of violent or non-violent collective action.

Policy support. We consider studies that measure support for policies and initiatives designed to improve the position the participants' ingroup in society.

Perceived injustice. We consider studies that measure perceptions that one is discriminated against because of one's group membership, that one's group faces discrimination in society, that one's group is relatively deprived compared to other groups, or that the deprivation and/or discrimination faced by one's group is unjust/illegitimate. We also consider expectations of fair treatment as the reverse of perceived injustice. We include studies that measure *personal* and/or *group* discrimination or a mixture of both.

3

2.1.5   Other criteria

We do not use any other criteria for eligibility for the meta-analysis.

2.2   Information sources

To find relevant studies, we retrieve records from electronic databases, including *Scopus*, an abstract and citation database of peer-reviewed literature, the *Scopus* secondary documents database[1], and *PsycINFO*, an abstract database of psychological literature. We also search the *ProQuest Dissertations and Theses* database for relevant grey literature. To find unpublished studies, we have sent a call for unpublished research to the mailing lists of the *European Association of Social Psychology* and the *Society for Personality and Social Psychology* (June/July, 2017). We send another call for unpublished research closer to publication in which we also include the mailing lists of the *International Society of Political Psychology*, the *Society for the Psychological Study of Social Issues*, and the *Society of Australasian Social Psychologists*. To achieve more complete coverage, we retrieve records that cite, or are cited by, the included studies. We also identify experts in the field and send a list of included studies to them, asking for any studies we have missed. We consider all relevant studies published before **April 1, 2019**, as well as all unpublished studies. We update our database searches **every four months** until we submit our manuscript for publication.

2.3   Search strategy

We developed a search strategy for locating studies that include the relevant predictor and outcome measures. We search titles, abstract, and keywords for relevant terms in four electronic databases. We search for relevant records in the *Scopus* database using the following search terms:

1. TITLE-ABS-KEY(*group AND (contact OR friendship) AND ("collective action" OR protest OR "collective behavio*r" OR "political behavio*r" OR "social change" OR "social justice"))

2. TITLE-ABS-KEY(*group AND (contact OR friendship) AND (policy OR policies OR "affirmative action" OR (politi* W/15 attitude*) OR (politi* W/15 preferenc*)) AND (redistribut* OR reparati* OR inequalit* OR equalit* OR injustice* OR justice* OR disadvantage* OR advantage* OR minorit* OR majorit*))

3. TITLE-ABS-KEY((contact OR friendship) AND ("perceived discrimination" OR "perceived *advantage" OR "relative deprivation" OR "group discrimination" OR "personal discrimination" OR "group deprivation" OR "perception* of discrimination" OR "perception* of group discrimination" OR "perception* of personal discrimination" OR "rac* discrimination"))

We use the same terms to retrieve relevant records from the *Scopus* secondary documents database. We retrieve relevant records from *PsycINFO* using similar search terms:

1. ("*group" AND (contact OR friendship) AND ("collective action" OR protest OR "collective behavio*r" OR "political behavio*r" OR "social change" OR "social justice")).ti,ab,id.

2. ("*group" AND (contact OR friendship) AND (policy OR policies OR "affirmative action" OR (politi* AND attitude*) OR (politi* AND preferenc*)) AND (redistribut* OR reparati* OR inequalit* OR equalit* OR injustice* OR justice* OR disadvantage* OR advantage* OR minorit* OR majorit*)).ti,ab,id.

---

[1]"A secondary document is a document that has been extracted from a Scopus document reference list but is not available directly in the Scopus database since it is not indexed by Scopus."

3. ((contact OR friendship) AND ("perceived discrimination" OR "perceived *advantage" OR "relative deprivation" OR "group discrimination" OR "personal discrimination" OR "group deprivation" OR "perception* of discrimination" OR "perception* of group discrimination" OR "perception* of personal discrimination" OR "rac* discrimination")).ti,ab,id.

We search for grey literature in the *ProQuest Dissertations and Theses* database using the following search terms:

1. NOFT(group AND (contact OR friendship) AND ("collective action" OR protest OR "collective behavio*r" OR "political behavio*r" OR "social change" OR "social justice"))

2. NOFT(group AND (contact OR friendship) AND (policy OR policies OR "affirmative action" OR (politi* N/15 attitude*) OR (politi* N/15 preferenc*)) AND (redistribut* OR reparati* OR inequalit* OR equalit* OR injustice* OR justice* OR disadvantage* OR advantage* OR minorit* OR majorit*))

3. NOFT((contact OR friendship) AND ("perceived discrimination" OR "perceived disadvantage" OR "perceived advantage" OR "relative deprivation" OR "group discrimination" OR "personal discrimination" OR "group deprivation" OR "perception* of discrimination" OR "perception* of group discrimination" OR "perception* of personal discrimination" OR "rac* discrimination"))

We export records from each database, compile meta-data (incl. authors, title, abstract, keywords, doi, link to full-text document), and remove duplicate records. If we find search terms to be inadequate, we will amend them and report these amendments. We use *Google Scholar* to find relevant records that cite any of the included studies.

2.4 Study selection

We select relevant studies in three steps. First, we screen records based on their title, abstract, and keywords. We divide records between two coders who each screen half of the records. We randomly select 100 records to be screened by both coders to calculate inter-rater agreement (Cohen's κ). For each record, coders decide whether the record meets the eligibility criteria (*yes*, *maybe*, *no*), or whether it is a relevant review article. We exclude all records for which coders answered "no", but keep all records for which coders answered "yes" or "maybe". Second, two coders each review each of the full-text manuscripts and code whether any sample in the manuscript fulfils the eligibility criteria (see Section 2.1). We calculate and report inter-rater agreement (Cohen's κ) for each eligibility criterion. Third, we resolve any disagreements between the two coders (by consensus) and exclude ineligible studies. At this stage, we also search for relevant records that cite, or are cited by, any of the included studies or relevant reviews. Figure 1 illustrates the study selection process. We use *Qualtrics* and/or *formr* (Arslan, Walther & Tata, 2018) surveys to screen and code studies.

2.5 Data collection

We collect the relevant data in two steps. First, we code moderating variables to be used in (exploratory) meta-regression analyses (see Section 2.7.3). We decide which moderators to code based on a first review of the included studies. At the very least, we code publication status (*published*, *unpublished*, *unpublished dissertation*), study setting (*ingroup*, *outgroup*, *location*), study design (*cross-sectional*, *longitudinal*, *quasi-experimental*, *experimental*), and study sample (*student*, *community*, *representative*). Two coders each review each of the included full-text manuscripts. We calculate and report inter-rater agreement (Cohen's κ) for each moderating variable.
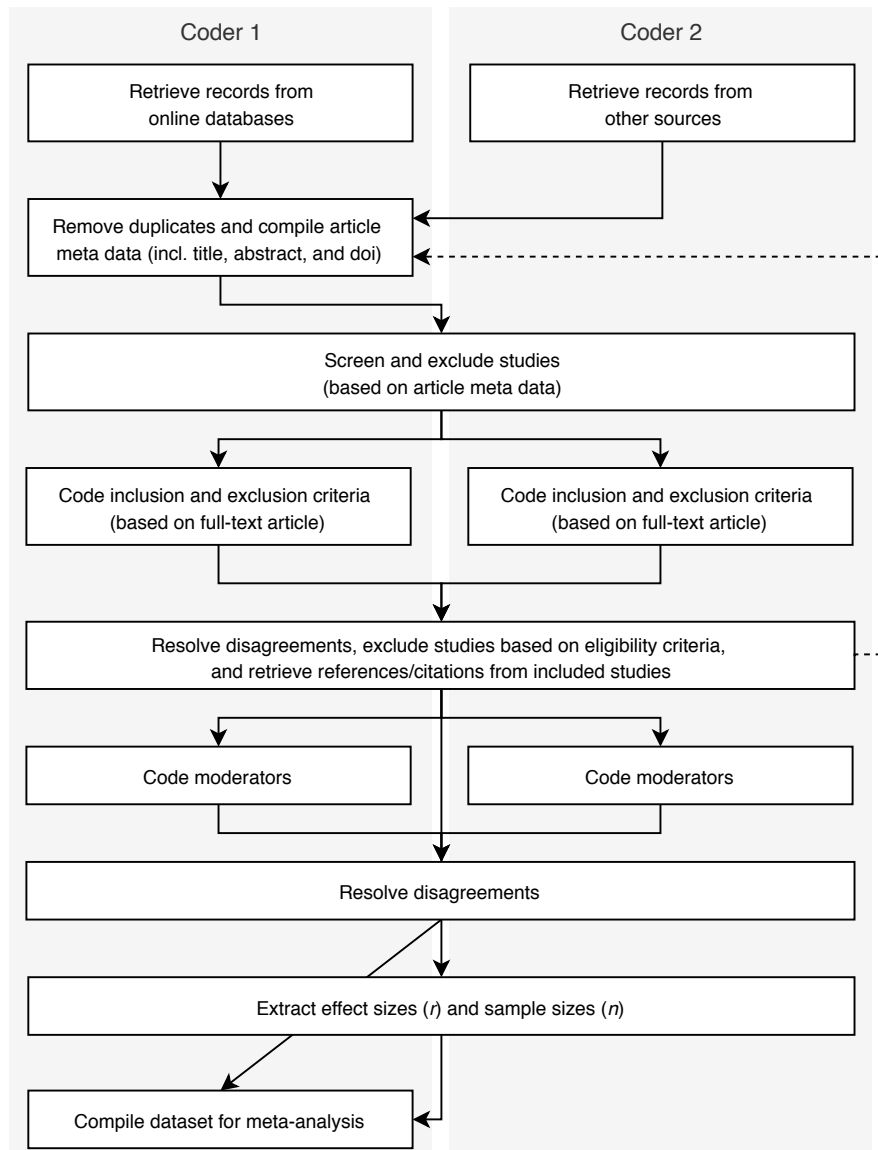
Figure 1. Diagram illustrating how two coders search for records, select studies, and collect data for this meta-analysis.

Second, we extract correlation coefficients (*r*) as the relevant measure of effect size, and extract sample sizes (*n*) to calculate standard errors for each sample's correlation coefficients (see Section 2.7). Where provided, we copy correlation coefficients from the relevant correlation matrices. Where other effect-size measure are provided, we convert the reported effect sizes to correlation coefficients using common conversion formulas (Borenstein, Hedges, Higgins & Rothstein, 2009). Where no effect sizes are provided, we attempt to contact the authors to obtain the relevant correlation matrices or other effect-size measures. If that fails, we impute correlation coefficients from beta coefficients (Peterson & Brown, 2005) if possible.[2] If we fail to obtain, calculate, or impute *r* or *n* for a sample, we report our failure to do so and exclude the sample from the meta-analysis. At this stage, we also look for samples that might have been reported more than once in the included literature. We consider author names, effect sizes, sample sizes, and other sample characteristics as indicators for duplicate studies. We might contact authors to resolve concerns about the identity of published studies.

We collect effect sizes for all predictor and outcome variables defined in Section 2.1, as well as all inter-scale correlations. Where more than one measure is reported, we extract effect sizes for all relevant measures. In addition, we extract effect sizes for five other measures that might be relevant for exploratory or future analyses: negative contact, ingroup contact, group identification, group-based anger, and outgroup attitudes.

2.6 Outcome selection

As discussed, we collect data for all relevant predictor and outcome variables defined in Section 2.1. When a study reports more than one measure of *intergroup contact* (Section 2.1.3), we prioritise the predictor variable in our confirmatory analyses (Section 2.7.1) that measure the most intense or intimate form of contact. We rate contact intensity according to the following ranking: friendship > quality of contact (positive contact) > quantity of contact > opportunity for contact. If it is unclear which of two predictor variables measures a more intense form of contact, we combine and average effect sizes across these variables. We do not consider negative contact or ingroup contact in our confirmatory analyses.

When a study reports more than one measure of one or more of the outcome variables defined in Section 2.1.4, we proceed as follows: If a manuscript's primary analysis focuses on one of several *collective action* measure, we prioritise that measure; otherwise, we combine and average effects sizes across outcome variables. If a study reports more than one measure of *policy support*, we combine and average effect sizes for all policies designed to improve the position of the participants' ingroup in society. If a study reports more than one measure of *perceived injustice*, we prioritise the measure closest to perceptions of injustice against the participants' ingroup (rather than against the participants themselves); if it is ambiguous which measure that is, we combine and average effect sizes across outcome variables.

Some study designs result in more than one effect size for the same (combination of) measures. If a *longitudinal* study reports results from more than two waves, we prioritise effect sizes spanning the inter-survey interval closest to one year. If a longitudinal study reports results from more than two waves and spans multiple years, we combine and average effects sizes from each one-year inter-survey interval. In our confirmatory analyses, we focus on the partial correlation between the relevant predictor variable (at time 1) and outcome variable (at time 2) controlling for initial levels of the

---

[2]$r = \beta + 0.05\lambda$ where $\lambda = 1$ when $\beta \geq 0$ and $\lambda = 0$ when $\beta < 0$.

outcome variable (cross-lagged correlation). If an *experimental* or *quasi-experimental* study reports comparisons between more than two conditions, we prioritise the effect size that compares two conditions that most closely resemble generic contact and no-contact conditions. If it is ambiguous which conditions those are, we combine and average effect sizes comparing experimental conditions to the control condition. We record all decisions about variable selection and data simplification in the published protocol.

2.7 Data synthesis (meta-analysis)

As discussed, we extract correlation coefficients (r) as the relevant measure of effect size. Correlation coefficients, however, are bounded between -1 and 1 and are not suitable for linear regression techniques. Therefore, we convert r to Fisher's z, which is unbounded and approximately normally distributed.

$$z = \frac{1}{2} \ln \left(\frac{1 + r}{1 - r} \right) \tag{1}$$

$$\sigma = \frac{1}{\sqrt{n - 3}} \tag{2}$$

$$r = \frac{\exp(2z) - 1}{\exp(2z) + 1} = \tanh(z) \tag{3}$$

For each sample, we use equation 1 to convert r to z, and use equation 2 to calculate the corresponding standard error σ from the sample size n. We use equation 3 to convert the effect-size estimates from our meta-analysis back to correlation coefficients.

We estimate Bayesian random-effects meta-analysis models in RStan (Stan Development Team, 2018) using z-transformed correlation coefficients as outcome variables:

$$z_{ij} \sim \text{Normal}(\theta_{ij}, \sigma_{ij}) \tag{Likelihood}$$

$$\theta_{ij} = \begin{cases} \mu + \mathbf{X}\beta_K + \beta_j \tau_J & \text{if } I_j = 1 \\ \mu + \mathbf{X}\beta_K + \beta_j \tau_J + \beta_i \tau_I & \text{if } I_j > 1 \end{cases} \tag{Regression}$$

where z_{ij} is the observed effect size in sample i of study j, σ_{ij} is the corresponding standard error, and θ_{ij} is the estimated effect size for sample i of study j. We estimate z_{ij} as a function of the estimated mean effect size μ and, if included, of a matrix $\mathbf{X}$ of K observed moderators multiplied by a corresponding vector of K coefficients β_K. We estimate two varying intercepts, β_i and β_j, with the corresponding standard deviations, τ_i and τ_j. For studies that contain only one sample ($I_j = 1$), we only estimate β_j, the study-wise deviation from the mean effect size μ. For studies that contain more than one sample ($I_j > 1$), we also estimate β_i, the sample-wise deviation from the study-specific effect size. We use the non-centred parameterisation for the varying effects (Betancourt & Girolami, 2015). The resulting model is a two-level random-effects meta-analysis.

Models assigned weakly informative prior distributions (Gelman, Simpson & Betancourt, 2017) to all estimated parameters:

$$\mu \sim \text{Normal}(0, 0.31605)$$
$$\beta_i \sim \text{Normal}(0, 1)$$
$$\beta_j \sim \text{Normal}(0, 1)$$
$$\beta_k \sim \text{Normal}(0, 1)$$
$$\tau_I \sim \text{Half-Cauchy}(0, 0.3)$$
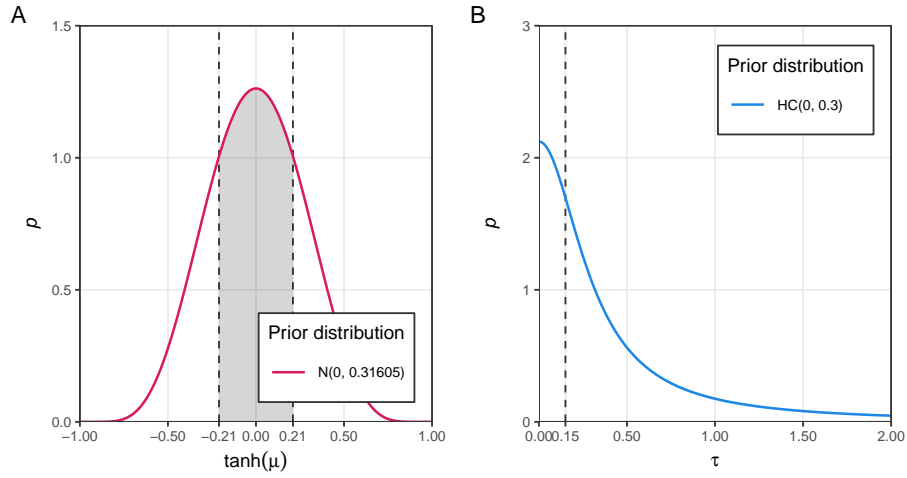$$\tau_J \sim \text{Half-Cauchy}(0, 0.3)$$

Figure 2. Prior distributions for μ and τ. (A) Prior distribution for μ converted from Fisher's z to r. The Normal$(0, 0.31605)$ distribution is centred around 0 and concentrates 50% of the most plausible values between $r = -0.21$ and $r = 0.21$. (B) Prior distribution for τ. The Half-Cauchy$(0, 0.3)$ distribution is strictly positive and concentrates 50% of the most plausible values below $\tau = 0.30$.

Figure 2 shows the prior distributions for μ, the mean effect size, and τ, the standard deviation of the varying intercepts. The prior distribution for μ is centred around 0 and concentrates 50% of the most plausible values between $r = -0.21$ and $r = 0.21$. We focus on $|r| = 0.21$ because it corresponds to the mean effect size observed in both Pettigrew and Tropp's (2006) meta-analysis of the contact literature and Richard, Bond Jr. and Stokes-Zoota's (2003) meta-analysis of effect sizes across a century of social-psychological research. The prior distribution for τ allocates 30% of plausible values below $\tau = 0.15$. We focus on $\tau = 0.15$ because it corresponds to the standard deviations observed in both Pettigrew and Tropp's (2006)[3] and Richard et al.'s (2003) meta-analyses. We chose this prior distribution following recommendations by Williams, Rast and Bürkner (2018).

2.7.1 Confirmatory analyses

We conduct confirmatory analyses to answer research questions Q1–Q3 (Section 1.2). Specifically, we run six two-level random-effect meta-analysis models (without moderators) to estimate the correlations between intergroup contact and perceived discrimination (r_1), policy support (r_2), and collective action (r_3), as well as the correlations between all three outcome variables (r_{12}, r_{13}, r_{23}), as illustrated in Figure 3. Each model includes only studies/samples that contain the relevant combination of variables. Each model considers predictor and outcome variables as defined in Section 2.6.

2.7.2 Sensitivity analyses

To assess the robustness of our conclusions, we conduct two kinds of sensitivity analysis. First, we assess to what extent our findings are sensitive to choosing different prior distributions. We repeat the main analyses (r_1, r_2, r_3) with narrower, $\mu \sim \text{Normal}(0, 0.1)$, and wider, $\mu \sim \text{Normal}(0, 1)$, prior distributions. Second, we assess to what extent our

---

[3]We estimated τ in a one-level random-effects meta-analysis using data from Pettigrew and Tropp (2006).
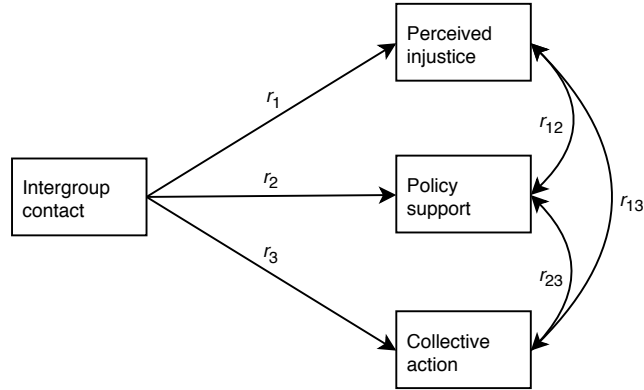
Figure 3. Path diagram illustrating all correlations that are estimated in our confirmatory analyses.

findings are sensitive to including or excluding influential studies. We rerun the main analyses J times—leaving out one of J studies each time—and record how much μ, the estimated mean effect size, varies across left-out studies. We calculate the mean absolute error (MAE) and compare the study-wise absolute errors ($|\mu_K - \mu_{-k}|$) to identify the most influential studies. We might report these analyses in an appendix to the published manuscript.

2.7.3 Exploratory analyses

We conduct exploratory analyses to estimate how much the strength of the relevant relationships varies across study designs, study settings, predictor variables, outcome variables, and other (potential) moderating variables. We explore these relationships in a series of two-level random-effect meta-analysis models with moderators. We enter moderating variables with K categories as $K - 1$ mean-centred dummy variables. We run additional meta-analysis models (without moderators) to estimate the partial correlations between the outcome variables and positive/negative contact and ingroup/outgroup contact—if we find enough studies that measure these predictor variables. We consider these analyses exploratory because they are guided by the available data.

2.8 Meta-biases

Publication bias, outcome reporting bias, and (other) questionable research practices in the relevant literature can lead meta-analyses to overestimate effect sizes. Developing methods to detect and correct for such biases is an active area of research, with no available method outperforming all others (Carter, Schönbrodt, Gervais & Hilgard, 2017). We assume that the literature underlying our meta-analysis is subject to these meta-biases. We accept that, in lieu of wide-spread adoption of preregistration, detecting and correcting for these biases is difficult or impossible. Nonetheless, we attempt to quantify meta-biases in the relevant literature as follows: First, we test for publication bias by comparing the mean effect sizes of published and unpublished studies. Second, we explore publication bias and other meta biases using statistical methods to detect and correct these biases. We select these methods based on the current state of the literature. We discuss meta-biases as a potential limitation in the published manuscript.

3   PRISMA-P (2015) checklist

Table 1. We explain how this protocol addresses each of the *preferred reporting items for systematic review and meta-analysis protocols* (PRISMA-P; Moher et al., 2015), and respond to items that we have not yet addressed. Shamseer et al. (2015) explain and elaborate all recommended items.

Section/Topic	#	Response
Administrative information		
Title:		
Identification	1a	The title identifies the report as the protocol for a meta-analysis.
Update	1b	This meta-analysis is not an update or extension of any existing systematic review or meta-analysis.
Registration	2	This protocol was preregistered with the Open Science Framework on the date stated on the title page.
Authors:		
Contact	3a	To facilitate blind peer review, we do not provide author names or contact information in this preregistered protocol. Below, we refer to the first and second author of the manuscript as A1 and A2.
Contributions	3b	A1 is the guarantor. A1 and A2 jointly wrote this protocol. A1 contributed the statistical methods to be used in the meta-analysis. A1 and A2 serve, respectively, as Coder 1 and Coder 2 (Figure 1). All authors approved this protocol.
Amendments	4	If we need to amend this protocol, we describe and justify all deviations from this preregistered protocol in the published protocol.
Support:		
Sources	5a	No funding has been received for this meta-analysis. We will disclose any future funding received for this meta-analysis.
Sponsor	5b	N/A
Role of sponsor or funder	5c	N/A
Introduction		
Rationale	6	Section 1.1 summarises the rationale for the meta-analysis.
Objectives	7	Section 1.2 outlines the research questions that the meta-analysis addresses.
Methods		
Eligibility criteria	8	Section 2.1 spells out all criteria for eligibility for the meta-analysis.
Information sources	9	Section 2.2 describes all intended information sources, as well as the planned dates of coverage.

Table 1. We explain how this protocol addresses each of the *preferred reporting items for systematic review and meta-analysis protocols* (PRISMA-P; Moher et al., 2015), and respond to items that we have not yet addressed. Shamseer et al. (2015) explain and elaborate all recommended items.

Section/Topic	#	Response
Search strategy	10	Section 2.3 describes our search strategy for four electronic databases, while the planned dates of coverage are described in Section 2.2.
Study records:		
Data management	11a	We do not use a dedicated systematic review data management software. We export records from electronic databases as .csv (or similar) files; we import .csv (or similar) files into *R* to compile meta-data and to remove duplicate records (Westgate, 2018). We use *Qualtrics* and/or *formr* (Arslan et al., 2018) surveys to screen and code studies (see Section 2.4).
Selection process	11b	Section 2.4 and Figure 1 describe how we select studies for inclusion in the meta-analysis.
Data collection process	11c	Section 2.5 and Figure 1 describe how we collect data for each included study/sample.
Data items	12	Section 2.5 lists all data that is collected; Section 2.6 explains all pre-planned data simplifications.
Outcomes and prioritization	13	Section 2.5 describes for which predictor and outcome variables we extract effect sizes; Section 2.6 explains which predictor and outcome variables are prioritised in our confirmatory analyses.
Risk of bias in individual studies	14	We explore study design as a potential source of bias (Section 2.7.3). We include all eligible studies in our confirmatory analyses, independent of their risk of bias.
Data synthesis	15a	We will conduct random-effects meta-analyses as described in Section 2.7.
	15b	Section 2.7 describes our method for combining data from studies as well as the planned summary methods (μ, τ_J, τ_I). Section 2.7.1 proposes confirmatory analyses for answering research questions Q1–Q3.
	15c	Sections 2.7.2 and 2.7.3 propose additional sensitivity and exploratory analyses.
	15d	Quantitative synthesis is appropriate for answering research questions Q1–Q3. We will discuss strengths and limitations of the available evidence in a narrative review.
Meta-bias(es)	16	Section 2.8 describes our approach to exploring meta-biases in the relevant literature.

Table 1. We explain how this protocol addresses each of the *preferred reporting items for systematic review and meta-analysis protocols* (PRISMA-P; Moher et al., 2015), and respond to items that we have not yet addressed. Shamseer et al. (2015) explain and elaborate all recommended items.

Section/Topic	#	Response
Confidence in cumulative evidence	17	We do not use a formal approach for assessing the quality of the available evidence, but discuss strengths and limitations of the available evidence in a narrative review.

References

Arslan, R. C., Walther, M. & Tata, C. (2018, 4th September). *formr: A study framework allowing for automated feedback generation and complex longitudinal experience sampling studies using R.* doi: 10.31234/osf.io/pjasu

Betancourt, M. & Girolami, M. (2015). Hamiltonian Monte Carlo for hierarchical models. In S. K. Upadhyay, U. Singh, D. K. Dey & A. Longanathan (Eds.), *Current trends in Bayesian methodology with applications* (pp. 79–101). Boca Raton, FL: CRC Press.

Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Chichester, UK: John Wiley & Sons.

Cakal, H., Hewstone, M., Schwär, G. & Heath, A. (2011). An investigation of the social identity model of collective action and the 'sedative' effect of intergroup contact among Black and White students in South Africa. *British Journal of Social Psychology*, *50*, 606–627. doi: 10.1111/j.2044-8309.2011.02075.x

Carter, E. C., Schönbrodt, F. D., Gervais, W. M. & Hilgard, J. (2017, 30th May). *Correcting for bias in psychology: A comparison of meta-analytic methods.* doi: 10.31234/osf.io/9h3nu

Dixon, J., Durrheim, K. & Tredoux, C. (2005). Beyond the optimal contact strategy: A reality check for the contact hypothesis. *American Psychologist*, *60*, 697–711. doi: 10.1037/0003-066X.60.7.697

Dixon, J., Durrheim, K. & Tredoux, C. (2007). Intergroup contact and attitudes toward the principle and practice of racial equality. *Psychological Science*, *18*, 867–872. doi: 10.1111/j.1467-9280.2007.01993.x

Dovidio, J. F., Love, A., Schellhaas, F. M. & Hewstone, M. (2017). Reducing intergroup bias through intergroup contact: Twenty years of progress and future directions. *Group Processes & Intergroup Relations*, *20*, 606–620. doi: 10.1177/1368430217712052

Gelman, A., Simpson, D. & Betancourt, M. (2017). The prior can often only be understood in the context of the likelihood. *Entropy*, *19*, 555. doi: 10.3390/e19100555

McKeown, S. & Dixon, J. (2017). The "contact hypothesis": Critical reflections and future directions. *Social and Personality Psychology Compass*, *11*, e12295. doi: 10.1111/spc3.12295

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., … Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*, 1–9. doi: 10.1186/2046-4053-4-1

Peterson, R. A. & Brown, S. P. (2005). On the use of beta coefficients in meta-analysis. *Journal of Applied Psychology*, *90*, 175–181. doi: 10.1037/0021-9010.90.1.175

Pettigrew, T. F. & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*, 751–783. doi: 10.1037/0022-3514.90.5.751

Reicher, S. (2007). Rethinking the paradigm of prejudice. *South African Journal of Psychology*, *37*, 820–834. doi: 10.1177/008124630703700410

Richard, F. D., Bond Jr., C. F. & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*, 331–363. doi: 10.1037/1089-2680.7.4.331

Saguy, T., Tausch, N., Dovidio, J. F. & Pratto, F. (2009). The irony of harmony: Intergroup contact can produce false expectations for equality. *Psychological Science*, *20*, 114–121. doi: 10.1111/j.1467-9280.2008.02261.x

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., … Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ*, *349*, g7647. doi: 10.1136/bmj.g7647

Stan Development Team. (2018). *RStan: The R interface to Stan*. Retrieved from `http://mc-stan.org/` (R package version 2.17.3)

Tropp, L. R. & Pettigrew, T. F. (2005). Relationships between intergroup contact and prejudice among minority and majority status groups. *Psychological Science*, *16*, 951–957. doi: 10.1111/j.1467-9280.2005.01643.x

van Zomeren, M., Postmes, T. & Spears, R. (2008). Toward an integrative social identity model of collective action: A quantitative research synthesis of three sociopsychological perspectives. *Psychological Bulletin*, *134*, 504–535. doi: 10.1037/00332909.134.4.504

Westgate, M. J. (2018). revtools: Bibliographic data visualization for evidence synthesis in R. *bioArXiv*. doi: 10.1101/262881

Williams, D. R., Rast, P. & Bürkner, P. (2018, 10th January). *Bayesian meta-analysis with weakly informative prior distributions.* doi: 10.31234/osf.io/7tbrm

Wright, S. C. (2001). Strategic collective action: Social psychology and social change. In R. Brown & S. L. Gaertner (Eds.), *Intergroup processes: Blackwell handbook of social psychology* (Vol. 4, pp. 409–430). Oxford, UK: Blackwell.

Wright, S. C. (2010). Collective action and social change. In J. F. Dovidio, M. Hewstone, P. Glick & V. M. Esses (Eds.), *The SAGE handbook of prejudice, stereotyping and discrimination* (pp. 345–360). London, UK: SAGE Publications.