

Has Intergroup Contact Delivered?

Annual Review of Economics

Matt Lowe

File structure

After unzipping, there should be four folders, with the following underlying folders/files:

1. code
 - a. `lowe_annrev_meta.do` – do-file for running almost all the analysis
 - b. `lowe_annrev_metaR.R` – R file for running the Bayesian analysis
2. data/raw
 - a. `1_lowe_contact_eligible_preregs.dta` – contains details of eligible pre-registrations, variables described below
 - b. `2_lowe_contact_studies.dta` – contains details of eligible papers/experiments, variables described below
 - c. `3_lowe_contact_effects.dta` – contains effect sizes and standard errors, along with other effect-level details, variables described below
 - d. `Combined_with_experimental.dta` – data used in the meta-analysis of Clochard (2024) for comparison, publicly available from <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DV/N/TRZUBI>
3. data/analysis
 - a. `4_lowe_contact_effects_explvl.dta` – contains aggregated experiment-level effects and standard errors, along with some other variables, fully described below. This dataset is outputted from the **collapse** switch in `lowe_annrev_meta.do`
 - b. `4_lowe_contact_effects_explvl.csv` – exported from the **collapse** switch in `lowe_annrev_meta.do`, this file is the key input into the Bayesian analysis in R.
 - c. `baggr_resultsX.csv` ($X=0, 12, 80$) – outputted from the Bayesian analysis in R. Used as an input for the **killerfig** switch in `lowe_annrev_meta.do`
4. figures
 - a. All figures created in `lowe_annrev_meta.do` are outputted here
5. tables
 - a. All tables created in `lowe_annrev_meta.do` are outputted here

Code steps

First change filepath at the top of `lowe_annrev_meta.do` and `lowe_annrev_metaR.R`, to direct to where you unzipped the replication files.

Then run `lowe_annrev_meta.do` switches **check_eligible** to **biascheck**. Then run `lowe_annrev_metaR.R` (for the robustness results, needs to be run three times: with `rho` at the top set to 0, 12, and 80). Then run the remaining switches in the do-file. The switches do the following:

- **check_eligible** – bring up key stats about eligible pre-registrations that are cited in the paper
- **check_papers** – bring up key stats about eligible papers/experiments that are cited in the paper
- **check_effects** – bring up key stats about treatment effects that are cited in the paper
- **collapse** – collapse effect-level dataset to experiment-level, outputted to `4_lowe_contact_effects_explvl.dta`
- **biascheck** – checks for p-hacking and related reporting biases, creates Figures 1 and A1
- **killerfig** – does random effects meta-analysis and reads in output of Bayesian analysis, creates Figure 2, Figures A2-A9, Figure A10 (with the local variable `rho` set to 0 at the top of the do-file), and Figure A11 (with `rho` set to 80)
- **genprob** – analysis of effects on generalized vs. non-generalized outcome variables using within-experiment variation, creates Table 2
- **bundVclean** – analysis of effects of bundled vs. clean treatments using within-experiment variation, creates Table A4
- **expfet** – analysis of predictors of experiment-level effects (e.g. whether contact is online or in-person), creates Table A2
- **globalsouth** – analysis of whether experiment in the Global South have different treatment effects, creates Table A3
- **onlineVoff** – creates Figure A12 comparing the effects of online with in-person contact
- **compare_clochard** – shows frequentist meta-analysis output when considering the subset of papers that overlap with Clochard (2024)

Datasets

Data on each pre-registration:

File name: 1_lowe_contact_eligible_preregs.dta

STATA dataset where each row corresponds to an eligible pre-registration. Summarizes the information in the pre-registration and contains the indicators that classify it as eligible. Includes registrations from AEA, EGAP, and OSF. The number of pre-registrations is higher than the number of studies, as some pre-registrations have not led to a paper, and a couple of studies have multiple pre-registrations.

Variables:

- `registry`: Registry site where we found the pre-registration. Either AEA, EGAP, or OSF.
- `preregID`: Unique ID for each pre-registration. Each registry site has a different format.
- `prereg_title`: Title of pre-registration. May differ from a study's title.
- `prereg_authors`: Authors of pre-registration.
- `prereg_year`: Year of initial pre-registration.
- `paper_available`: Dummy variable to indicate whether a paper for the pre-registered study is available. We searched for papers using Google search, author's websites, and through emails to authors requesting a draft.
- `prereg_before_analy`: Dummy variable to indicate whether the study was pre-registered before the data was analyzed (required for eligibility).
- `rand_intergroup`: Dummy variable indicating whether the study has a randomized intergroup contact treatment (required for eligibility).
- `prej_outcome`: Dummy variable indicating whether the pre-registration has at least one outcome related to prejudice and/or intergroup relations (required for eligibility).
- `publish_journal`: Journal title if the pre-registration corresponds to a published study. NA if not published.
- `publish_year`: Year of publication if the pre-registration corresponds to a published study. NA if not published.
- `description`: Description or abstract of the study as shown in the pre-registration. If not available in a specific description/abstract field in the pre-registration or PAP, it includes other sections of the pre-registration that summarize what the study is about.
- `prereg_notes`: Miscellaneous notes about the pre-registration.
- `prereg_other_title`: Alternative title for the study that came out of the pre-registration.
- `duplicate`: Dummy variable indicating whether the pre-registration corresponds to a study for which we have an earlier pre-registration.

Data on each paper:

File name: 2_lowe_contact_studies.dta

STATA dataset where each row is a study that corresponds to one or multiple eligible pre-registrations. Provides details of registration, publication, and characteristics of the study. Each study contains one or two pre-registered experiments; in the case of two, we list the relevant information for each experiment.

- `registry`: Registry in which the experiment(s) of the study was pre-registered. Either AEA, EGAP, or OSF.
- `paperID`: Unique ID created to identify a study within a registry.
- `date_first_reg`: Date of first pre-registration.
- `paper_authors`: Authors of the study.
- `paper_title`: Title of the study.
- `paper_year`: Year of latest draft we consulted.
- `published`: Dummy variable to indicate whether the study is published.
- `journal`: Journal of publication. NA if not published.
- `reg_before_intervention`: Dummy indicating whether the experiment was pre-registered before the intervention began. For studies from AEA pre-registrations.
- `registered_when`: Timing of pre-registration (e.g. prior to any research activities). For studies from EGAP and OSF pre-registrations.
- `prereg_url`: URL for the pre registration(s) associated with the study.
- `prereg_ID`: Pre-registration ID for AEA pre-registrations.
- `treat_highVlow`: Dummy variable indicating whether the study has high intergroup contact (treatment condition) versus low intergroup contact (control condition), holding social interaction fixed.
- `treat_someVno`: Dummy variable indicating whether the study has some intergroup contact (treatment condition) versus no intergroup contact (control condition), holding social interaction fixed.
- `treat_ingroupVctrl`: Dummy variable indicating whether the study has ingroup contact treatment versus pure control treatment (i.e. no contact at all).
- `treat_outgroupVctrl`: Dummy variable indicating whether the study has outgroup contact treatment versus pure control treatment (i.e. no contact at all).
- `treat_horse_race`: Dummy variable indicating whether the study can horse-race contact treatment against another type of treatment.
- `contact_type`: Dummy variable indicating whether the intergroup contact intervention in the study requires participants to interact in person, online with video, or online without video.
- `online_contact`: Dummy variable indicating whether participants interact online (with or without video).
- `more_than_one_type`: Dummy variable indicating whether the study has more than one type of intergroup contact intervention.
- `allport_equal_status`: Dummy variable indicating whether the intergroup contact intervention involves equal status between participants.

- `allport_coop`: Dummy variable indicating whether the intergroup contact intervention involves cooperation between participants.
- `allport_comm_goals`: Dummy variable indicating whether the intergroup contact intervention involves participants having a common goal.
- `allport_support`: Dummy variable indicating whether the intergroup contact intervention involves support from authority.
- `allport_conditions`: Number of Allport conditions satisfied (0 to 4).
- `SD_outcome_available`: Dummy variable indicating whether the standard deviations are available in the paper or data for each pre-registered outcome.
- `SD_ctrl_available`: Dummy variable indicating whether the standard deviations for the control group are available in the paper or data for each pre-registered outcome.
- `Nexperiment1`: Number of individuals in the experiment. Includes all participants across all treatment arms.
- `Nexperiment2`: Number of individuals in the second experiment. Applicable to studies with two pre-registered experiments.
- `female_share1`: Percent of female participants in the experiment. -99 if unknown.
- `female_share2`: Percent of female participants in the second experiment. -99 if unknown. Applicable to studies with two pre-registered experiments.
- `mean_age_1`: Mean age of participants in the experiment. In some cases, it includes only a subset of participants in the relevant analyses sample (e.g. analyses done for one of the groups in the contact intervention).
- `mean_age_1_details`: Details on how mean age of participants is obtained.
- `mean_age_2`: Mean age of participants in the second experiment, if applicable. In some cases, it includes only a subset of participants in the relevant analyses sample (e.g. analyses done for one of the groups in the contact intervention).
- `mean_age_2_details`: Details on how mean age of participants in the second experiment is obtained.
- `contact_group1`: Relevant group identity for intergroup contact in the experiment (e.g. religion: contact between members of different religions).
- `contact_group2`: Relevant group identity for intergroup contact in the second experiment. Applicable to studies with two pre-registered experiments.
- `country`: Country where the experiment(s) took place.
- `pap_public1`: Dummy variable indicating whether the pre-registration for the experiment includes a publicly available pre-analysis plan.
- `pap_public2`: Dummy variable indicating whether the pre-registration for the second experiment includes a publicly available pre-analysis plan. Applicable to studies with two pre-registered experiments.
- `data_available`: Dummy variable indicating whether the data obtained from the experiment is available to the public.
- `data_url`: URL to obtain the experiment's data, if available.
- `paper_notes`: Miscellaneous paper notes.

- `global_south`: Dummy variable indicating whether the experiment took place in a country in the Global South.
- `paper_authors_year`: Lists authors and year of the paper.
- `in_clochard`: Dummy variable indicating whether the study is included in Clochard's 2024 meta-analysis.

Data on each pre-registered outcome in each study:

File name: 3_lowe_contact_effects.dta

STATA dataset where each row corresponds to an effect for a pre-registered outcome in an experiment. Summarizes the characteristics of the outcome, the comparison being drawn to obtain the effect, and the statistics for the effect. Each pre-registered outcome may have one or more effects, depending on the timepoints at which the outcome was measured, the number of comparisons made, and whether the effects are reported separately for each group.

Variables:

- `registry`: Registry in which the experiment(s) of the study was pre-registered. Either AEA, EGAP, or OSF.
- `paperID`: Unique ID created to identify a study within a registry.
- `registry_paperID`: Unique ID created to identify each study across registries.
- `experimentID`: Unique ID of experiment within a paper (most papers have only one).
- `registry_expID`: Unique ID of experiment across all studies.
- `reg_exp_comp_grpID`: Unique ID to identify each group x comparison in each experiment (e.g. compare high vs low outgroup contact for local participants).
- `reg_out_grpID`: Unique ID to identify each outcome x group in each experiment (e.g. measure Bias among locals).
- `effectID`: ID for outcome-timing combination within this experiment.
- `comparisonID`: ID for comparison being made within this experiment (sometimes multiple in an experiment).
- `groupID`: ID of group within this experiment for which we have the effect (sometimes reported separately).
- `paper_authors`: Authors of the study.
- `paper_year`: Year of latest draft we consulted.
- `paper_authors_year`: Lists authors and year of the paper.
- `paper_title`: Title of the study.
- `prereg_url`: URL for the pre registration(s) associated with the study.
- `primary_prereg_outcome`: Description of this pre-registered primary outcome.
- `reg_before_analysis`: Dummy variable indicating whether the outcome was pre-registered before analysis.
- `collapseID`: ID for effects to be collapsed with independence. Differs rarely from `effectID`. Only differs in cases where we want to collapse two or more effect sizes on different endpoints assuming independence. E.g. a study with a treatment effect on

some outcome Y1 for Hindus and an effect for some outcome Y2 for Muslims – these are two different endpoints, but since they are estimated for non-overlapping populations, we collapse assuming independence. Hence these two effects would have the same `collapseID`.

- `group_name`: Name of group when effects are reported separately by group.
- `outcome_type1`: Outcome type, either behavioral measure, explicit evaluation, political and cultural attitudes, indirect measures, or unrelated (mostly following Paluck et al. 2019 coding). Behavioral measures include behaviors towards intergroup relations, hypothetical or not, even if not directly towards the outgroup. We also include third party reports on behavior. Explicit evaluations include explicit attitudes, feelings, and beliefs about the outgroup. Political and cultural attitudes include attitudes towards topics related to intergroup relations or commonly associated with prejudice. Indirect measures include implicit attitude tests, the evaluation of hypothetical vignettes, list experiment, and other measurements that avoid social desirability bias. Unrelated refers to outcomes that are not about intergroup relations or prejudice. An outcome can have multiple outcome types.
- `outcome_type2`: Outcome type, either behavioral measure, explicit evaluation, political and cultural attitudes, indirect measures, or unrelated. An outcome can have multiple outcome types.
- `outcome_type3`: Outcome type, either behavioral measure, explicit evaluation, political and cultural attitudes, indirect measures, or unrelated. An outcome can have multiple outcome types.
- `outcome_type4`: Outcome type, either behavioral measure, explicit evaluation, political and cultural attitudes, indirect measures, or unrelated. An outcome can have multiple outcome types.
- `outcome_behav`: Dummy variable indicating that the outcome is classified as a behavioral measure.
- `outcome_explic`: Dummy variable indicating that the outcome is classified as an explicit evaluation.
- `outcome_indire`: Dummy variable indicating that the outcome is classified as an indirect measure.
- `outcome_poli`: Dummy variable indicating that the outcome is classified as a political and cultural attitude.
- `outcome_unrelated`: Dummy variable indicating that the outcome is classified as unrelated to intergroup relations or prejudice.
- `specific_general`: Indicates whether the outcome measures behaviors/attitudes/beliefs towards the outgroup in general (not including the specific people met) or whether it includes the specific people met.
- `generalized`: Dummy variable indicating whether the measure doesn't include specific people met (=1 if it doesn't include, =0 if it does).
- `behavioral_incent`: Indicates whether the outcome measures a real or incentivized behavior (=1) versus not (=0). Self-reports of own behavior or intended behaviors are not

considered real or incentivized. Variable equals 0.5 when an outcome is a mix of components that are real behaviors and components that are not (e.g. index outcome).

- `outcome_in_paper`: Dummy variable indicating whether the outcome is reported in the paper. Some outcomes are pre-registered but then do not appear in the paper.
- `treatment_comparison`: Comparison drawn between the treatment and the control group. Either high versus low intergroup contact, some versus no intergroup contact, contact with outgroup bundled with other component(s) of the intervention versus pure control, or contact with ingroup bundled with other component(s) of the intervention versus pure control.
- `treatment_comparison_notes`: More details on relevant treatment comparison.
- `breadth`: How broad is the contact intervention—number of additional outgroup members met for this comparison. -99 if missing.
- `breadth_details`: More details on breadth.
- `contact_duration_best_guess`: Best guess for duration of contact in hours.
- `contact_duration_literal`: Literal description of contact duration in paper.
- `days_until_measurement`: Days in between contact ending and measurement for this effect. 0 if measured on the same day.
- `in_main_paper`: Dummy variable indicating whether this treatment effect is reported in the main paper, as opposed to the appendix. If the treatment effect is reported in a figure and we need to obtain the exact estimates elsewhere, it is still considered in the main paper.
- `in_abstract`: Dummy variable indicating whether this treatment effect is reported in the abstract, and if so, reported accurately (e.g. an effect reported as a significant change shows indeed a statistically significant change).
- `estimate`: Point estimate for this treatment effect. Coded such that more positive is associated with improved intergroup relations.
- `se`: Standard error for the point estimate.
- `sample_size`: Sample size in regression which includes this treatment effect.
- `num_clusters`: Number of clusters in regression which includes this treatment effect.
- `sample_size_relevant_groups`: Sample size for the relevant treatment groups which identify this effect. -99 if missing.
- `num_clusters_relevant_groups`: Number of clusters for the relevant treatment groups which identify this effect. -99 if missing.
- `control_sd`: Control group standard deviation for this outcome. -99 if missing.
- `full_sample_sd`: Full sample standard deviation for this outcome. -99 if missing.
- `immediate`: Dummy variable indicating whether `days_until_measurement` is 0.
- `months_until_measurement`: Months in between contact ending and measurement for this effect. `days_until_measurement` divided by 30.
- `best_sd`: Control group standard deviation if available, otherwise full sample standard deviation, otherwise -99 if both standard deviations are missing.
- `effect_size`: Standardized effect size. Point estimate divided by `best_sd` if non-missing.

- `se_std`: Standardized standard error. Standard error divided by `best_sd` if both standard error and `best_sd` non-missing.
- `var_std`: Standardized variance. `se_std` squared.
- `t_stat`: Absolute value of point estimate divided by standard error.
- `bundled_out`: Dummy variable indicating whether comparison is outgroup bundled with other component(s) of the intervention versus pure control group.
- `bundled_in`: Dummy variable indicating whether comparison is ingroup bundled with other component(s) of the intervention versus pure control group.
- `contact_type`: Dummy variable indicating whether the intergroup contact intervention in the study requires participants to interact in person, online with video, or online without video.
- `online_contact`: Dummy variable indicating whether participants interact online (with or without video).
- `allport_equal_status`: Dummy variable indicating whether the intergroup contact intervention involves equal status between participants.
- `allport_coop`: Dummy variable indicating whether the intergroup contact intervention involves cooperation between participants.
- `allport_comm_goals`: Dummy variable indicating whether the intergroup contact intervention involves participants having a common goal.
- `allport_support`: Dummy variable indicating whether the intergroup contact intervention involves support from authority.
- `allport_conditions`: Number of Allport conditions satisfied (0 to 4).
- `global_south`: Dummy variable indicating whether the experiment took place in a country in the Global South.

Experiment-level effects dataset:

File name: 4_lowe_contact_effects_explvl.dta

STATA dataset where each row corresponds to a pre-registered experiment. Dataset includes experiment-level frequentist aggregated treatment effects, for different sets of underlying treatment effects (e.g. considering only clean effects of contact), and for different assumptions of the within-experiment effect size correlation (ρ).

Variables:

- `registry`: Registry in which the experiment(s) of the study was pre-registered. Either AEA, EGAP, or OSF.
- `paperID`: Unique ID created to identify a study within a registry.
- `registry_paperID`: Unique ID created to identify each study across registries.
- `experimentID`: Unique ID of experiment within a paper (most papers have only one).
- `registry_expID`: Unique ID of experiment across all studies.
- `paper_authors`: Authors of the study.
- `paper_year`: Year of latest draft we consulted.

- `paper_authors_year`: Lists authors and year of the paper.
- `contact_duration_best_guess`: Best guess for duration of contact in hours.
- `online_contact`: Dummy variable indicating whether participants interact online (with or without video).

The variables below have a suffix from s1-s8, denoting different eligibility criteria for underlying treatment effects. The eight set of criteria map to the eight effects shown in Figure 2. The criteria are: (1) clean effects of contact, (2) clean effects on generalized outcomes, (3) clean effects on generalized outcomes, only in-person contact with all Allport conditions met, (4) same as (3) but keeping only interventions with four or more hours of contact. (5)-(8) are the same, but keeping only effects of bundled contact interventions.

- `behavioral_incent_mean_s1` – experiment-level mean of `behavioral_incent` for the effects that satisfy the criteria.
- `generalized_mean_s1` – experiment-level mean of `generalized` for the effects that satisfy the criteria.
- `Days_until_meas_mean_s1` – experiment-level mean of `days_until_measurement` for the effects that satisfy the criteria.
- `exp_effect_size_s1` – experiment-level standardized effect size
- `exp_se_stdX_s1` – experiment-level standardized standard error, assuming that $\rho=0.X$, for $X=0, 10, 12, 15, 20, 40, 60, 80$